

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:32:00 ; Search time 10277.7 Seconds  
(without alignments)  
4505.438 Million cell updates/sec

Title: US-09-513-888-1  
Perfect score: 9048  
Sequence: 1 gcctttccagaccctgcc.....tgcattttctcacgcccctct 9048

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues  
Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_om: \*  
4: gb\_ov: \*  
5: gb\_ph: \*  
6: gb\_pl1: \*  
7: gb\_pl2: \*  
8: gb\_pl3: \*  
9: gb\_pr2: \*  
10: gb\_pr3: \*  
11: gb\_ro: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: em\_fun: \*  
15: em\_hum1: \*  
16: em\_hum2: \*  
17: em\_in: \*  
18: em\_om: \*  
19: em\_or: \*  
20: em\_ov: \*  
21: em\_pat: \*  
22: em\_ph: \*  
23: em\_pl: \*  
24: em\_ro: \*  
25: em\_sts: \*  
26: em\_sy: \*  
27: em\_un: \*  
28: em\_vl: \*  
29: gb\_ba3: \*  
30: gb\_in1: \*  
31: gb\_in2: \*  
32: gb\_in3: \*  
33: gb\_pl3: \*  
34: gb\_pr4: \*  
35: em\_ba1: \*  
36: em\_ba2: \*  
37: em\_htg1: \*  
38: em\_htg2: \*  
39: em\_htg3: \*  
40: em\_htg4: \*  
41: em\_htg5: \*  
42: em\_htg6: \*  
43: em\_htg7: \*

44: em\_htg8: \*  
45: em\_htg9: \*  
46: em\_htg10: \*  
47: em\_hum3: \*  
48: em\_hum4: \*  
49: em\_hum5: \*  
50: em\_hum6: \*  
51: gb\_pr5: \*  
52: gb\_pr6: \*  
53: gb\_pr7: \*  
54: gb\_htg1: \*  
55: gb\_htg2: \*  
56: gb\_htg3: \*  
57: gb\_htg4: \*  
58: gb\_htg5: \*  
59: gb\_htg6: \*  
60: gb\_htg7: \*  
61: gb\_htg8: \*  
62: gb\_htg9: \*  
63: gb\_htg10: \*  
64: gb\_htg11: \*  
65: gb\_htg12: \*  
66: gb\_htg13: \*  
67: gb\_htg14: \*  
68: gb\_htg15: \*  
69: gb\_htg16: \*  
70: gb\_htg17: \*  
71: gb\_htg18: \*  
72: gb\_htg19: \*  
73: gb\_htg20: \*  
74: gb\_htg21: \*  
75: gb\_htg22: \*  
76: gb\_htg23: \*  
77: gb\_sts1: \*  
78: gb\_sts2: \*  
79: gb\_vil: \*  
80: gb\_vil2: \*  
81: gb\_pat1: \*  
82: gb\_pat2: \*  
83: em\_hum0: \*  
84: gb\_htg24: \*  
85: gb\_pr8: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	9046.8	100.0	9108	10	AF123653	AF123653 Homo sapi	
c	2	4926.8	54.5	172162	64	AC025853	AC025853 Homo sapi
	3	4199.8	46.4	5492	10	AF123659	AF123659 Homo sapi
	4	807.2	8.9	1515	10	AF123656	AF123656 Homo sapi
	5	807.2	8.9	1614	10	AF123655	AF123655 Homo sapi
	6	807.2	8.9	1692	10	AF123657	AF123657 Homo sapi
	7	807.2	8.9	1722	10	AF123658	AF123658 Homo sapi
	8	346	3.8	633	10	AF123654	AF123654 Homo sapi
	9	229.2	2.5	163683	64	AC025412	AC025412 Homo sapi
c	10	226.2	2.5	152511	60	AC019269	AC019269 Homo sapi
	11	225.6	2.5	179538	63	AC024740	AC024740 Homo sapi
	12	224.6	2.5	159440	34	AL355392	AL355392 Human DNA
	13	224.2	2.5	177560	59	AC016695	AC016695 Homo sapi
	14	224	2.5	227949	65	AC026803	AC026803 Homo sapi
	15	223.6	2.5	160659	66	AC036235	AC036235 Homo sapi
	16	223.6	2.5	180516	84	CNS01DSH	AL121784 Homo sapi
	17	223.6	2.5	190128	54	AC007374	AC007374 Homo sapi
c	18	223.6	2.5	215945	74	AL355337	AL355337 Homo sapi
	19	223.4	2.5	89818	8	AC002126	AC002126 Homo sapi
c	20	223.2	2.5	141292	72	AF235106	AF235106 Homo sapi
c	21	223	2.5	156345	62	AC023015	AC023015 Homo sapi

22 223 2.5 203088 73 AL162505 Homo sapi  
c 23 222.6 2.5 189623 60 AC018640 Homo sapi  
c 24 222.6 2.5 189663 55 AC009244 Homo sapi  
c 25 222.4 2.5 170664 74 AL358013 Homo sapi  
c 26 222.2 2.5 191162 9 AC008865 Homo sapi  
c 27 222 2.5 162176 72 AL137780 Homo sapi  
c 28 222 2.5 169236 69 AC073167 Homo sapi  
c 29 222 2.5 251664 55 AC008908 Homo sapi  
c 30 221.8 2.5 110000 55 AC008749\_0 Homo sapi  
c 31 221.8 2.5 154621 75 AL391836 Homo sapi  
c 32 221.6 2.4 184316 63 AC024022 Homo sapi  
c 33 221.2 2.4 188740 63 AC024045 Homo sapi  
c 34 221 2.4 175839 8 AC004491 Homo sapi  
c 35 220.6 2.4 170866 64 AC026203 Homo sapi  
c 36 220.6 2.4 189268 63 AC024021 Homo sapi  
c 37 220.4 2.4 151217 64 AC026164 Homo sapi  
c 38 220.4 2.4 160095 59 AC018458 Homo sapi  
c 39 220 2.4 158859 66 AC004895 Homo sapi  
c 40 220 2.4 191979 54 AC005867 Homo sapi  
c 41 219.8 2.4 208202 57 AC013446 Homo sapi  
c 42 219.6 2.4 180638 72 AL157709 Homo sapi  
c 43 219.6 2.4 278310 69 AC069592 Homo sapi  
c 44 219.2 2.4 144293 62 AC022408 Homo sapi  
c 45 219.2 2.4 156895 57 AC012652 Homo sapi

ALIGNMENTS

RESULT 1  
AF123653 AF123653 9108 bp DNA PRI  
LOCUS Homo sapiens FEZ1 (FEZ1) gene, complete cds.  
DEFINITION AF123653  
VERSION AF123653.1 GI:4572463  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9108)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,  
and its expression is altered in multiple human tumors  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
MEDLINE 99199287  
REFERENCE 2 (bases 1 to 9108)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE Direct Submission  
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
source  
1. 9108  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="9p22"  
/chromosome="8"  
mRNA join(<112..456,1707..2510,4912..>5553)  
/gene="FEZ1"  
/product="FEZ1"  
gene <112..>5553  
/gene="FEZ1"  
CDS join(112..456,1707..2510,4912..5553)  
/gene="FEZ1"  
/codon\_start=1  
/product="FEZ1"  
protein\_id="AAD23834.1"  
/db\_xref="GI:4572464"  
/translation="MGVSVLISGHSFHKCRASQYKLRKSHLKLNRYSGLLRF  
FPSQDGHGKSSSKMKSDEFFYIKVSKARGSHHPDYTALSSGLDGGAGVDPDPT  
FPKLMPFSNQLMESEKGAVRPTAFKPVLPFRSGAILHSSPESASHLHPAPPKPKQ

|||||  
Db 901 CTGCCCTGGTGCATTAGCTCCTTTCAAGCTGGCTGGGGCATCTTTGAATACAG 960  
QY 901 gctggagttttatccattaccagagactaggcaagagagcccaagcagcactgagaa 960  
Db 961 GCTGGAGTTTGTCAATTCATTTACCAGAGACTAGGGCAAGAGAGGCCACGACTGAGAA 1020  
QY 961 atccagccctcacacagctcaagccctcgctgctccacgagtgagacactgaaatcaat 1020  
Db 1021 ATCCAGCCCTCACACAGCTCAAGCCCTCGTGGTCCACAGAGTGGACACTGAATCAAT 1080  
QY 1021 ttctctattcagtcctctgcccctggccctgggggaaatgaaatcccccggcttgattact 1080  
Db 1081 TTTCTCTATTTCAGTCTCTGCCCTTGCCCTGGGGAAATGAATCCCGGGCTTGAATTAAT 1140  
QY 1081 aggaagagcccttatgtttcatagagcattcagcttttcaaatlaagggcgttgtaa 1140  
Db 1141 AGGAAGAGCCCTCTATGTTTGCATAGACATTCAGCTTTTCAAAATTAAGGGCTTTGTA 1200  
QY 1141 actgtgaagcactctaccagggaataattacagttttaaaaaaggatcgtgattggagtg 1200  
Db 1201 ACTGTGAACCACTTACCAGGGAAATTAACAGTTTAAAAAAGGATCGTGAITGGAGTG 1260  
QY 1201 agctcccaacccctgtaaggagccaggtcctgtgtccttgcctccagggccttaatggaag 1260  
Db 1261 AGCCTCCCAACCTGTAAAGAGGCCAGGTCCGCTGTCTTGTCTCCAGGCTTAATGGAAGAG 1320  
QY 1261 gcaagtgaacaggaaggaatggaacctaaagagagacagcaagctcgccagcctgatg 1320  
Db 1321 GCAGTGAACAGGAAGAGGATGGACCTTAAGAGGGACAGCAAGCTCGGCCAGCCTGTATG 1380  
QY 1321 ccttaacttgcccacacagagaccttagagcagagcctcaagatggtatttatcaacctc 1380  
Db 1381 CCTTAATTGCCCCACACAGAGACCTTAGAGAGGAGCCCTCAAGATGGTATTTATCACCTC 1440  
QY 1381 gggagggtcggggaagctggtggcaggttgctatttcatagaacaaagtgcoccaaagtcg 1440  
Db 1441 GGGAGGGCTGGGCAAGCTGGTGGCAGGTGCTATTTCATAGAACAAGTGCCCAAGTCG 1500  
QY 1441 ccattagggttttccctcctaagagagatgacattcagctgcttcaaaagcaacaggcaa 1500  
Db 1501 CCATTAGGGTTTTTCCTCCTTAAGAGAGATGACATTCAGCTGCTTCAAGCAACAGGCAA 1560  
QY 1501 ggtcgtctgagacaattgaccaagagggtgctgctgcgctcagagagcccaagactggc 1560  
Db 1561 GGTCTGCTGAGACAATTGACCAAGAGGGTGTGCTGCTGCTCAGAGACCCACAGACTGCC 1620  
QY 1561 tcaagtcgagcgcgtgctcctggggaggaggtgcaatgcgcgcgagggagggcatga 1620  
Db 1621 TCAAGGTCGGCAGCGCTGCTGGGGAGGGGTGCAATGCGCGCGCAGGGGAGGCATGA 1680  
QY 1621 gtcaacgggtcctttctctataagggctcagagaggtgcagtgagggcccaacagcct 1680  
Db 1681 GTCACCGGGTCTTTTCTCTACAGGGTCCGAGAAGGGTGCAAGTGAGGCCCCACAGCCT 1740  
QY 1681 tcaagcctgtgctgcacaggtcaaggacatcctgcactcctcccgagagtgccagcc 1740  
Db 1741 TCAAGGCTGTGCTGCCACGGTCAAGAGACCATCTCTGCACATCTCTCCCGGAGAGTGCCAGCC 1800  
QY 1741 accagctgaccccgccctccagacaagcccaagagagcaggtgagcctggcctgt 1800  
Db 1801 ACCAGCTGCACCCGCCCTTCCAGACAAGCCCAAGAGCAGGAGCTGAAGCCTGGCCCTGT 1860  
QY 1801 gctctggggcgtctcagactccggccggaactccatgctccagcctgcccacacagca 1860  
Db 1861 GCTCTGGGGCGGTGTGAGACTCCGGCCGGAACCTCCATGTCTCCAGCCTGCCACACAGCA 1920  
QY 1861 ccagagagcctaccagctggaccgcgtgttcacacccgtgggaccaccaagccgttttg 1920  
Db 1921 CCAGCAGCAGCTACCAGCTGGACCCGCTGGTCACACCCCTGGGACCCCAAGCCGTTTTG 1980  
QY 1921 ggggctcccgcccaacaatcaaccagggaatcgtctccagagacagcagcatgatgagcc 1980  
|||||

Db 1981 GGGGCTCGCCCCACAACTACCCAGGGCATCTCTCTCCAGCAGCAACATGATGAGCC 2040  
QY 1981 tgaagctctgtctcttctccgagcaggtagcaagctggccactcgacaagaggcagaca 2040  
Db 2041 TGAAGGCTGTCTCTTCTCCGACGAGGTAGCAAGCTGGGCCACTCGAACAAGAGGACACA 2100  
QY 2041 agggccctcgtgtgtcgcgtcccccatctccacgagcaggtgcagcatcccaggagctgg 2100  
Db 2101 AGGGCCCTCGTGTGTCTGCCCTCCCATCTCCACGAGCAGTGCAGCATCCAGGAGCTGG 2160  
QY 2101 aacagaagctgttgagagggaggcgccctcagaagctgcagcgcagcttttgaggaga 2160  
Db 2161 AACAGAAAGCTGTGGAGAGGGAGGGCGCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGA 2220  
QY 2161 agagcttctccacagcctgctctacgagagcgccgcgcgtgcagggcagcagctgg 2220  
Db 2221 AGAGGCTTGCCCTCCAGCCTTGGCTTACGAGAGCGCGCGCGCTGCAGGAGAGAGCTGG 2280  
QY 2221 agggcccgaggcccaaaaggcgcaacaagctcaagcaggccctcgagaagagagccagcg 2280  
Db 2281 AGGGCCCGAGCCCAAAAGCGGCAACAAGCTCAAGCAGGCCCTCGCAGAAGAGCAGCGCG 2340  
QY 2281 cgcagcaggtcctgcacctgcaggttactgcagcttcagcagagagaagcgagcctccgc 2340  
Db 2341 CGCAGCAGGTCTCGCACTTGCAGGTACTGCACTTTCAGCAGGAGAAAGCGGCACTCCGC 2400  
QY 2341 agagctcgcagcctcatgaaggagcagacgtcgtgcagaccagctcaggtctcctacg 2400  
Db 2401 AGAGGCTCCAGAGCCCTCATGAAGGAGCAGACCTGCTGGAGACCAAGCTCAGGTCCTACG 2460  
QY 2401 agaggagaagaccagcttcgcccgcgctgcagagaccagctggggaggtgagggccac 2460  
Db 2461 AGAGGGAGAAGACCAGCTTTCGGCCCCGCTGGAGGAGACCCAGTGGAGGTGAGGCCAC 2520  
QY 2461 acagggtcctatgggttctgggtggttcagcggtttgcgccagtaacccccctctctctgg 2520  
Db 2521 ACAGGGCTCATGGGTTGGGTGCTCAGCGGTTTGGCGCCAGTACCCCTCTCTCTCTGG 2580  
QY 2521 tctggccaatagctgcgaacacagaccgcagcagcagcagcagcagcagcagcagcagc 2580  
Db 2581 TGTGTCCTCAATAGCTGCAACACAGACCGCGCAGCAAGCGGGCTTAATGCTGGCTT 2640  
QY 2581 tatccccaaaaggggctccctgcaaacctagtgtgggggatacgcattacatctagct 2640  
Db 2641 TATCACCCAAAGAGGGGCTCCTCTGCAACCATGTTGGGGGATCGACTTACATCTGAGCT 2700  
QY 2641 tctcctgtccccaccatcacctcatggtcctcagtttcccaagtgcagccat 2700  
Db 2701 TCTCTGTCTCCACCATCACCTCATGGCTCTAGATTTTCAGTTTCCCAAGTGAAGCCAT 2760  
QY 2701 taaatcatgaagccggaagccagatgaccaagcccgagccagcagcagcagcagcagcagc 2760  
Db 2761 TAAATCATGAAGCCGGAAGCCAGATGACCAAGCCCGCAGCAGGCTGTGGGCTGACCTCCC 2820  
QY 2761 ttccatcagctccagagaggtcagaagaagaacacgcgtgcctgagttcagggcggggc 2820  
Db 2821 TTCCATCAGCTCCCAGGAGCTCAGAAGAAGAACAGCCGTGCCTGAGTTCAGCGCGGGC 2880  
QY 2821 caggggcccaagagacagaaatgcatttggcttggaggaggagcagcagcagcagcagc 2880  
Db 2881 CAGGGCCCCAAGAGAGACAGAAATGATTTGTGCTTTGGAGGGAGGAGCTGCACCCACT 2940  
QY 2881 agtaagagggagccctatttgggtggcaggtttcagtgatgaagtgggcagcagcagcagc 2940  
Db 2941 AGTAAGAGGGAGCCCTATTGTTGGTGGCAGGTTTCAGTGAATGGAAGTGGCCACTCTCTTGTG 3000  
QY 2941 ggtgaagtgaacttctatttgggtgagctgagatgaaacccctagggagaggaagaaagag 3000  
Db 3001 GGTGAAGTGAACCTTCTATTGTTGGTGGCTGAGATGGAATACCTTAGGAGAGAGAAAGAG 3060  
QY 3001 tccccactcacacacttacacactcacacactcacactcacactcacactcacactcacact 3060  
Db 3061 TCCCCACTCACACACTTACACACTCACACACTCACACTCACCTCACCCGGCTCACAGCTGGA 3120

QY 3061 gaggcatctgtacctgacggtctgtgagaaacccataaacctctgcatctattagtgga 3120  
Db 3121 GAGGCATCTGTACCTGACCGTCTGGAGAACCCCATACCTCTGCATCTATTAGTGGGA 3180  
QY 3121 agcagctttctcacagcctggtggtctggatgactcatgagttcaagcccatcggtg 3180  
Db 3181 AGCAGCTTTTCTCACAGCCTGGTGGTCTGGATGACTCATGGAGTTCAAGGCCATCGTTG 3240  
QY 3181 aggccttttaactgtcgcaacccagcttggctgtctcagtgccctgacctccccagtt 3240  
Db 3241 AGGCTCTTTTACATGTTCGCACCCAGCTTGGTCTGTCCAGTGCCTGCCCTCACCCCACT 3300  
QY 3241 cagagtcctaaactcagtcctacaccaaaccctggtctatgtgcaagtcaacaacccagtg 3300  
Db 3301 CAGAGTCCAAATCTCAGTCTACACGAAACCCCTGGCTATGTGCAAGTCAACAACAGTG 3360  
QY 3301 gtttaacttgccactgctggcagctgtatcaccccaatttaacaccaatggtattggt 3360  
Db 3361 GTTTAACTTGCCACTGCTGGCAGCTGATACACCCCATTTAAACACCAATGTTGGTT 3420  
QY 3361 ttggtgtcagcctgatttctgcatcgatgtttatgccacatctctgacctccccct 3420  
Db 3421 TTGGGTGTACGCTGATTTCTGTCTCATCGATGTTTATGCCACATCTCTGACCTCACCCCT 3480  
QY 3421 gcatgcacccagccctccctctctctgctactgagtaaaagactacctcacaattcac 3480  
Db 3481 GCATGCACCCAGCCCTCTCTCTCTGCTACTGAGTAAAGACTACCTCACAAATTCAC 3540  
QY 3481 tgcgtaccagtgactagtcactgctggtcttgatgcagagcccaatccacatctgtc 3540  
Db 3541 TGCTGTACCCAGTACTAGTATCATCTGGCTTGGATGCAGAGCCCAATCCACATCTGTC 3600  
QY 3541 aaacgaggaatcattttcc 3600  
Db 3601 AAACGAGGAATCATTTCC 3660  
QY 3601 ccatacaaatgtggcgaagcaatgataaaaacggaagccacaggttagaccatgtgt 3660  
Db 3661 CCATCAAAATTTGGCCAGAGCAATGATGAAACCCGAAGCCACAGTTAGACCCATGTGT 3720  
QY 3661 ctctggatcttgccactctggggtcatgggagaccaaagccagctgctgcatcttaag 3720  
Db 3721 CTCTGGATCTTGGCCATCTGGGTCATGGGAGACCAAGGCCAGTCTGGCTGAATCTTAAG 3780  
QY 3721 agtgaatgaagtcceagagcatgtgctctacagaatgagattcttggaactagcctggaag 3780  
Db 3781 AGTGAATGAAGTCCAGAGCATGTGGCTCTACAGAAATGGATTCTTGGAACTAGCCCTGGAAG 3840  
QY 3781 ccaccttcacatttctcttcagtagaaaattcccttgcctcagtgaaacactgcac 3840  
Db 3841 CCACCTTCACATTTCTTTACAGTAGAAAATTTCCCTCTGGCCCTCAGTGAACACTGAC 3900  
QY 3841 agtctgtgagaaaatccgacctaccacccaggatgctgtctggagaccaaagaatttcattcc 3900  
Db 3901 AGTCTGGAGAAAATCCGACCCTACCCAGGATGCGTGTGGGACCAAGAATTTTCATTCC 3960  
QY 3901 aaggccaacctgtattcatgccagaaggagtgacacagtcatggtgaggaatgggc 3960  
Db 3961 AAGGCCAACCCCTGTATTATGCGACGAAGGGAGTGACACAGTCACTGGGTGAGGCATGGGC 4020  
QY 3961 ctgggtttgaacctcagcttgaccacttatgatccagtgattgttaaatacatattagccat 4020  
Db 4021 CTGGCTTTGAACCTCAGCTTGACCACTTATGATCCAGGTGATTGTAATACATTAACCAT 4080  
QY 4021 ggtggcaatggggtatagttaaactgttgggatacaactctctactcttatactttat 4080  
Db 4081 GGTGGCAATGGGGTATAGTGATTAACCTGTTGGGATCAAACTCTCTACTCTTATACTTTAT 4140  
QY 4081 attttatatatatataataatatatatatatattagacctcaggtggtcaactcac 4140  
Db 4141 ATTTTATATATATATATAATATATATATATATATATATATAGCCCTCAGGCTGGTCACTTAC 4200

QY 4141 cagctgtttgtctatcataacccctctctgtccctcagttctcatgtatgataaattgaggacta 4200  
Db 4201 CAGCTGTTTGGCTATCATACCTCTCTGTGCCCTCAGTTTCATTGATGATAAATTGAGACTA 4260  
QY 4201 ctaatagtagctacttctatcggtgttaaggaaatagatgagcaaatgtatggtcttggcac 4260  
Db 4261 CTATAGTAGCTACTTCTATCGGGTGTAAAGGAATAGATGAGCAAAATGTATGCTTGGCAC 4320  
QY 4261 ttaataaactactcaaaattatttagtgaagtatgtttataataataactctctgttggct 4320  
Db 4321 TTAATAAACAATAAATATTAGTGAAGTATGTTTATAATAATATACTTCTGTGTGGCT 4380  
QY 4321 aggcgtggtggtcagcctcgaatccacagactttggagcagagcagcagcagcagac 4380  
Db 4381 AGGCTGTGGTGGCTCAGCCTGCAATCCCACTCTTGGAGGAGCAGAGGCAGGCAGACAC 4440  
QY 4381 ttgaggtcaggaaatcagatcagcctggtcccaactgaggaacccctctctactaaaa 4440  
Db 4441 TTGAGGTCAGGAATTCGAGATCAGCTGGCCAACATGAGNAACCCCGTCTCTACTAAA 4500  
QY 4441 atacaaaaatcagcaggcatggtgaggtgtctgttaatccccactacttggaggtg 4500  
Db 4501 ATACAAAATCAGCCAGGCATGTTGGCAGGTCTCTGTAATCCAGCTACTTGGAGGCTG 4560  
QY 4501 aggcaggagaatcagagggagcgaggttgcagtgagccaagatcacgccactacac 4560  
Db 4561 AGCAGGAGAAATCAGAGGGAGCGGAGGTTCAGTGAGCCCAAGATCACGCCACTACACC 4620  
QY 4561 ccagcctagggtgacaaagcagacttctcaaatataaacaataataactatgtgtc 4620  
Db 4621 CCAGCCTAGGTGACAAAGCGAGCTTCAAAATATTAAACAATAATATATATATGTGTC 4680  
QY 4621 attatacatgatattatttttatttatttatttatttatttatttatttatttatttattt 4680  
Db 4681 ATTATACATGATGATTTATTTATTTTATCTATTTATAGCTTAGCTAGCTGATAAAGCTGG 4740  
QY 4681 araaaggtcacagcaatgtcagcttacttccagattgacaaaagctggaatgctaac 4740  
Db 4741 ARAAAGGTACAGCAATGTTACGCTTACTTTTCAGATTGGACAAAGCTTGAATGCTTAAC 4800  
QY 4741 accggggccacgcgcattcccgagtggtgtttatttttagcagctgagctgtcacttccct 4800  
Db 4801 ACCGGCCACCGCATCCGAGTGGCTTGGTTATTTTAGSCAGCTGAGCTGCATCTCCCT 4860  
QY 4801 ggttaagagacactcactcttggcactctgtctcaccccccctcggcaggtgtgcaag 4860  
Db 4861 GGGTAAGGACACTCACTCTTGGCCTCTGTCTCCACCCCTCGGCAGGTGTGCCAG 4920  
QY 4861 aagtcagcagatctccctcctgaagcagcagctgaagagatccacagcagagtgaaac 4920  
Db 4921 AAGTCAGGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGTCCACAGCGAGGTGAAC 4980  
QY 4921 gcaagggttagcgagatcctggtctcaaggcacagctgaaggaacacgcgggccaagctg 4980  
Db 4981 GCCAAGGCTAGCGAGATCCTGGTCTCAAGGCACAGCTGAAGCACACGGGGGCAAGCTG 5040  
QY 4981 gaggccctggagctgagaccaggaacctggagggcgccctgcgcaacaaggccctggag 5040  
Db 5041 GAGGCCCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTCGCCACCAAGGCCCTGGAG 5100  
QY 5041 ctggaggtctgtgagaatgagctcagcgcaagaagaacagagcgagctgctgcggag 5100  
Db 5101 CTGGAGGTCTGTGAGAAATGAGCTGCAGCGCAAGAAACGAGGCGGAGCTGCTGCGGGAG 5160  
QY 5101 aaggtgaacctgtgtgagcagagctgcagagctgcggggccagggccctggtcccgcc 5160  
Db 5161 AAGGTGAACCTGCTGGACGAGAGCTGCAGGAGCTGCGGGCCAGGCCCTGCCCCGC 5220  
QY 5161 gacatggggcgcccaacttccccagagacgtccctgcctgcagcgagagctggagcgg 5220  
Db 5221 GACATGGGGCGGCCCACTTCCCCAGGACGCTCCCTGCCCCCTGCAGCGGAGCTGAGCGG 5280  
QY 5221 ctgcgggcccagctgcgggagagcggcaaggccatgaccagatgtcctcgggcttccag 5280

|||||  
Db 5281 CTGCGGCGAGCTGCGGAGAGCGGCAAGGCCATGACAGATGTCTCTCGGCTTCCAG 5340  
QY 5281 catgagcgctcgttggaagagagagagaaagtgattcagttaccagaaacagctg 5340  
Db 5341 CATGAGCGGCTCGTGTGAAGAGAGAGAGAGAGAGGTGATTACGTACCAGAAACAGCTG 5400  
QY 5341 cagcagagctacgtggccatgtaccagcggaaccagcgctcggaagagccctgcagcag 5400  
Db 5401 CAGCAGAGCTACGTGGCCATGTACCAGCGGAACACAGCGCTGGAGAAGGCCCTGCAGCAG 5460  
QY 5401 ctggcaagtggaagacagcccggaagcccttggaggttgacctggaaggggctgacatc 5460  
Db 5461 CTGCGCAGTGGGACAGCGCCGGGAGCGCTTGGAGTTGACCTGGAAGGGCTGACATC 5520  
QY 5461 cctcagagacatcatagcactagatctgagggctgctggaaagcgagctcgtgg 5520  
Db 5521 CCTACGAGGACATCATGCCACTGAGATCTGAGGGGCTGCGCTGGGAAGCGGAGCTGGG 5580  
QY 5521 gacctggcactggagcgaggctctccggtgcatcccccctgctcagcaattcagaccc 5580  
Db 5581 GACCTGGCACTGGAGGCAAGGCTCTCCGCTGCATCCCCCTGCTCAGCAATTACAGACCC 5640  
QY 5581 ctctgagagacccaactccctgggaacagacagccaggaaccccgaggggagggagga tg 5640  
Db 5641 CTCTGAGAGACGCCACTCCCTGGGACACAGACCAGGACCCCGAGGGAGGGCAGGATG 5700  
QY 5641 gaccttccctctctctgagtccagtgctccagtgctcaccagccctgcagccaccagacgtca 5700  
Db 5701 GCGTTTCCTTCCTCTCTGATGTCCAGTGCCTCACCAGCCCTGCAGCCACCAGACGTCA 5760  
QY 5701 ggcctgactctcttgcttcccagagatgggtccagggctctgctgttggttaa 5760  
Db 5761 GCGCTGACTCTCTGGCTTTCCAGGAGATGGGTCCAGGGTCTGCTCTGCTTTGTTAA 5820  
QY 5761 gggctccctaaacttggccttggcttcgaata tagatatcctctccccctccctccagggaa 5820  
Db 5821 GGGCTCCCTAAACTTTGGCTTTGTCGTAATAGATATCTCTCCCTCTCCAGGGAA 5880  
QY 5821 ggtggccacagaagaagcggtccctccgctctctctcattcccaacctcttctctcc 5880  
Db 5881 GGTGGCCACAGAAAGACAGCGGCTCCCTCCGCTTCTCATCCCAACCTCTTTTCTCTCC 5940  
QY 5881 tggacacattgaaatccttggaaatagaagaagcacaatatagcagaagccttgga 5940  
Db 5941 TGGACACATTGGATGCCCTTGGAAATAGAAAGAGCCATATATGACCAGAGCCTTGGAA 6000  
QY 5941 ccagcccatcagaacctgagctatttctctctggccgagaggtgtaggggtggaatga 6000  
Db 6001 CCAGCCCATCAGAACCTCAGCTATTTTCTCTGGCCGACAGAGGTGTAGGGTGGAAATGA 6060  
QY 6001 gcgcggggaagctggcttgaacctcaggctgtccagcccggaagccacaggaa 6060  
Db 6061 GCCCGGGGAAGCTGGCTTTGAAACCTCAGGGGTGTCCAGAGCCCGGCAAGCCACAGGAA 6120  
QY 6061 ggaaggagagacagcagccagcagtgatgagacctgcacacccagagagagcag 6120  
Db 6121 GGAGGGAGAGACAGCGCCACCCAGCAGTGTGAGAGCCCTGCCACACAGAGAGGGCAG 6180  
QY 6121 agggagaatccaagggttgagagccagtggcggtgtagtgccagccctggggccagcc 6180  
Db 6181 AGGGAGAATCCAAGGCTTGAGAGCCAGTGGCGGTGATGGCCAGCCCTTGGGCCCCAGCC 6240  
QY 6181 cctgttacttggtcttgaaatggagctgagcagcctctggagccagtgacatttg 6240  
Db 6241 CCTGTTTACTGTGTTCTGAAAATGGGAGCTGAGCAGCCCTCTGGACAGCCAGTGACCTTGG 6300  
QY 6241 acctcgtgaccactcttcttaagccatagacctgagccctgagcctggctgctggga 6300  
Db 6301 ACCTCGGTGACCACCTCTTCTTTAAGCCATAGACCCCTGAGGCCCTTGGGCTGGGTGGGA 6360  
QY 6301 agggaggggtgaaacacccgtgaaccagaggggtgtggcttcttcagkacccctcagggagc 6360  
|||||

Db 6361 AGGGAGGTTGAAACCAACCGTGAACCAAGAGGCTGTGGCTTTCCAGKCAACCCCTCAGGAGC 6420  
QY 6361 ctccccatctgtccagctggggccagagagctggagtcctcctacctgcttcaagttggccg 6420  
Db 6421 CTCCCATCTGTCCAGCTGGGGCCAGAGGCTGGGACTCCCTACCTGCTTCAGCTTGGCCG 6480  
QY 6421 gggctactctggaatgttttccctccccagaa tcaagcttttcttgatccagaagag 6480  
Db 6481 GCGGCTACTCTGGAATGTTTTCCCTCCCCAGAAATCAAGCTTTTGTGTTGATCAGAAAG 6540  
QY 6481 cccatatcaactaagatgcatatatgtatctgggcatcttctcctcctctctcagca 6540  
Db 6541 CCATATCACTAAGATGGCATATATGTGATCTGGGCATTTTCTCTCTGCTGTACAGCCA 6600  
QY 6541 ggttagcggaaccccttcccccttagcacctcagggctgagttcttggtttcttagag 6600  
Db 6601 GGTTTAGCGGCAAACTTTCCCTTTAGCACTTACCACTTCAAGGCTGAGTTCTGGGTTC 6660  
QY 6601 gtccagagcggtcctcctcagagcgccaggaagccagagcccaagcaggaagaaagagc 6660  
Db 6661 GTCAGGACGGCTCTCAGAGCGCCAGGAAGCCAGAGCCCAAGCAGAGGAGGAAAGAGC 6720  
QY 6661 atacacacagcagtgatgaatagcctggccacagcagcactcctcctccactcaagacccc 6720  
Db 6721 ATACACACAGCAGTGTGAATAGCTGGCCACCAGCCATCTCTCCCTCAGCTCAAGACCC 6780  
QY 6721 catttgcacactaaagatccagagagcagctcccttctcagagcttggcagtg 6780  
Db 6781 CATTTGTCCSAGACTAAAGATCCAGAGAGCAGCTCCCTTTCTCAGAGCTTGGGCAGTG 6840  
QY 6781 cccagggaggtccagggtttctctcagatgtcgagcgagggcggtggtagagagag 6840  
Db 6841 CCCAGGGAGTCAGGGTTCTCTGCAGATGTGCGAGGCGGAGCGGTGGTAGAGAG 6900  
QY 6841 ataaaggtyggagttctctctgttgggtcagggattttattttattttatagagac 6900  
Db 6901 ATAAAGGTGGAGTTTCTCTGTGTGTGGTTTTCAGGGATTTTATTTTAAATTTATGAGAC 6960  
QY 6901 agggcttctctctccccagcgtggagtcagtgagcagctccttctcagagcttggcagcc 6960  
Db 6961 AGGGTCTTCTCTGTCCCCAGGCTGGAGTGGATGGCATGGCATGATAGCTACATGTCAGCC 7020  
QY 6961 tcatactcctgggtcagaacaaatcctcctgcctcagccttccaaactagctgggactacag 7020  
Db 7021 TCATACTCTCGSCTCAAGCAATCTCTCTGCTCAGCTCAGCTTCCAACCTAGCTGGACTACAG 7080  
QY 7021 gtgcgcacacogtgcctggctgtaacttttcttctttttgtaggggaggggtctgtttt 7080  
Db 7081 GTGCGGCCACCGTGGCTGCTAACTTTTCACTTTTGTAGGGACGGGCTCTCGTTTT 7140  
QY 7081 gttgcaaaagctgggtctcaaaacttggcctcagaacaaatcccaactgccttggcctccaa 7140  
Db 7141 GTTGCCAAAGCTGGTCTAAACTTTGGGCTCAAGCAATCCCACTGCTTGGCCTCCCAA 7200  
QY 7141 agtgctgagatgcatgagcagcctgctgctggccagattttcttcttcttcttcttct 7200  
Db 7201 AGTGCTGATTTGCAGATGTGAGCCACCGTGCCTGCCAGATTTTCTTTTATCTTCTT 7260  
QY 7201 tcttttcttttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 7260  
Db 7261 TCTTTTCTTTTGTGCTTTGCTTTTTCATTTTCTTTTGTAGGGACGGGCTCTCGTTTT 7320  
QY 7261 agttctattttttagcttagccacagctgtgttctcagggacagcactcccttcccaat 7320  
Db 7321 ATGTCTATTTTGTAGCTTAGCCACAGCTGCTTCTCAGGACAGCATCCCTTCCACAT 7380  
QY 7321 gctgcgcctgctgcctgctgagatgagggagcgtcttggaacttgcagtgctcag 7380  
Db 7381 GCCTGGCCCTGCTGCTGCTGAGATGAGGAGGGAGCGCTTGGGAACCTTGGGAGTCCAAG 7440  
QY 7381 gccagtccttcttctctctcctcagcctggtggtccttagagaccccgaggtgaggggtgg 7440  
Db 7441 GCCAGTCCCATTTCTGCTCTGCTCAGCGCTGGCCCTTAGAGACCCCGAGGTAGGGGTGG 7500

Qy 7441 ggagatgctctctctcttgccccccgcctcatgggtctcctagccttcccttaagtgcggg 7500  
Db 7501 GGAGATGCTTCTCTCTGTCGCCCGCCCGCCCTCAAGGCTCTAGCCCTTCCCTGAGTGCGGG 7560  
Qy 7501 ctgagggcagagtcaccttttctgtggtggctctaccttccctgctccctgaggttaaacg 7560  
Db 7561 CTGAGGCCAGAGTCACCTTTTCTGTGGCTGGCTCTACCTTCTCTGCTGCCCTGAGGTTAAACG 7620  
Qy 7561 gtgccatctcctcatctcaaacacagagagaggttttcttggaatttcaaaacattgct 7620  
Db 7621 GTGCCCATCTCTGCCATCTCTCAACACACAGAGAGGCTTTTCTGGAATTTCAAAACCATTTGCT 7680  
Qy 7621 cttagtcccaagctaggcttaaacctgaatctacaagccaaaagtccctccctcctga 7680  
Db 7681 CTTAGTCCCAAGCTAGGCTTAACCTGGAATCTACAGCCAAAAGTCCCTCCCTGCCCTGA 7740  
Qy 7681 gggcagttaccttccattgggacagctccagacccaagtcacaagatgccccattccttgcg 7740  
Db 7741 GGGCAGTACCTTCCATTTGGGCACAGTCACAGACCCCAAGTCAAAGATGCCCATTTCCCTGGC 7800  
Qy 7741 cctcagcctcagttccttcatttccaccagggcgtgctgttctgttgagtttttctcccca 7800  
Db 7801 CCTCAGCCCTCAAGTCTCTTCAATTTCCACAGGCCGTGGCTTCTTTGAGTTTTCCTCCCA 7860  
Qy 7801 gtgagactgccccacggagacagagaaaggctgctgctccctcccccagggctggagacc 7860  
Db 7861 GTGAGACTGCCCCACGGAGACAGAGAAAGGCTGGCTGCCCTCCCCAGGCTGGAGACC 7920  
Qy 7861 ccccccaactccaggaaagcagtcagtcagtcagtcgtctgctgcacagctgtgctgag 7920  
Db 7921 CCCCCCAACTCCAGGAAAGCAGTCAGAGTCCAGTGTCTTGCCCTCAGAGGTTGCCCTGAG 7980  
Qy 7921 aagaagtggtgtccacacccaggggaagccctcagggcggagggctgtgctccggcctagt 7980  
Db 7981 AAGAAGTGGCTGCCACCCACAGGGGAAGGCCCTGAGGGGAGGCTGTGTCTCGCCCATGGT 8040  
Qy 7981 gtcccggtaccttccatcacagagagtgagagcttctccatatctccatggtccctgtc 8040  
Db 8041 GTCCCGGTACTTCTTCATACACAGAGAGTGACAGCCCTTCTCCATATCTCATGCGCCCTGTC 8100  
Qy 8041 ccaggccggccagatgtgtccccccagggcctgtctcctcagtcacaaagtggcagatgc 8100  
Db 8101 CCAGGCGGGCCAGATGTGTCTCCCGCCAGGCCCTTGCTACGTCCAAGGTGCGAGATGTC 8160  
Qy 8101 ttccctgggtgtccacacagcccccccccagagtggtccacacogtggcactagaatgcaag 8160  
Db 8161 TTCCCTGGGCTGCCACCCAGCCCGCCCGCCAGAGTGGCCACCGTGGCCTAGAAATGCAAG 8220  
Qy 8161 tatectgcgaccttgcaacctcaaccttctgtgggtgttcttctcctgacctgtccaaaag 8220  
Db 8221 TATCTGGGACCTTGCACACCTTCACTTCTGTGGGTGTCTTCTCTGCGCTGTCCAAAAG 8280  
Qy 8221 cgccctactattcttgaccatgcagatgctcctcctcctcctggaagaggctcttgacag 8280  
Db 8281 CGCCCTCACTATTCTTGGACCATGCCAGATTCTGCTCTCTCTGGAAGAGGCTCTGGACAG 8340  
Qy 8281 cagaagcctccaagcacagagcctgtgccccagggccccacagaggggtgggttctcctgcct 8340  
Db 8341 CAGAAGCCTCCAAGCACAGAGCCTGGCCCCAGGCCGCCAGACAGGGGTGGGCTTCTCTGCCCT 8400  
Qy 8341 tccctctgggacgcctcctgtggcagaccactgaacctcctcctcctcctcctcctcctcct 8400  
Db 8401 TCCCTCTGGGACGCTCTGTGGCGACCCACTGACCCCACTGGATGGACCAACCTGCTCT 8460  
Qy 8401 gtccccaaagacgcctcagagagagcagcactcccatccatccactcaccacagatcgga 8460  
Db 8461 GTCCCCAAAGAGCGCTCGAGGAGAGACGACGACTTCCGCATCACCCTCACCAAGGATCCGA 8520  
Qy 8461 ctctcctcctggacctgggaacgactgactgtcactcaggggttccctcctcctcctcctcctcct 8520  
Db 8521 CTCTGCCCTTGACCTGGGAACGACTGAGCTGTCTACGGGGTTCCTCTCTAGTCTCTCCACG 8580

Qy 8521 tgaactctgccaggcacacacagccccctatatagcactagctcactcagctcagctgggactggatat 8580  
Db 8581 TGAACTCTGCCAGGCACACAGCCCCCTATAGCACTGAGCTCAGCTGAGACTGGGATAT 8640  
Qy 8581 gggggcatctcttccccagagaggcactcagtgagcctcctctgtgctcctgccccagctctgg 8640  
Db 8641 GGGGCATCTCTTCCCCAGAGAGGCACCTCAGTGAGCCTCCTGTGCTTGCCCTGGCCCCAGCTGG 8700  
Qy 8641 gccatctctttaggttagacagttgccccaaactaagcagggcctggtggaagagcagca 8700  
Db 8701 GCCATCTCTTAGTGAGACAGTTGCCCGAAACTAAGACCAGGCTGGCTGGAGGACGACGA 8760  
Qy 8701 gcttgggagagggatttccctgcagactcaagcctcatcgcgtggtgctgcacatga 8760  
Db 8761 CTTGGGAGAGGGATTTCCTTCGACAGCTCAAGCCATCATCGGTGGTGTGTCATGA 8820  
Qy 8761 cagaggtgcacccctggggcagcgggctgctcaccacacctcttctgtcgaaggtggcctt 8820  
Db 8821 CAGAGGCTGCACCCCTGGGCCAGCGGGCTGCTCACCCACCTCTTGTGCAAGGTGSCCTT 8880  
Qy 8821 tgtgtgcgctgcaggcagagctggagccccccacagagcagggcctgggacggaccagc 8880  
Db 8881 TGTGTGCGCTGTCAGGACAGCTGGAGCCCCCAGCAGAGGCGAGGCTGGGACGACCAGC 8940  
Qy 8881 atctggaagatgacatagttattttcttcttcttcttcttcttcttcttcttcttcttct 8940  
Db 8941 ATCTGGAAGATGACATAGTTATTTTCTCTTTGTGGTTCCTGTGTTGTTGTTGTTGTTGCT 9000  
Qy 8941 tttagacagcttcattttattttgacgtcacttttggccatgttggccatgtaaactattttggcaa 9000  
Db 9001 TTTGACAGCTTCATTTTATTTTGTGACGTCACTTTTGGCCATGTAAACTATTTTGTGGCAA 9060  
Qy 9001 ttttatgttttatttatgaataaagaatgccatttctcagccctct 9048  
Db 9061 TTTTATGTTTATTTATTAATAAAGATGCCATTTCTCACGCCCTCT 9108

RESULT 2  
AC025853/c  
LOCUS AC025853 172162 bp DNA HTG 26-MAY-2000  
DEFINITION Homo sapiens chromosome 8 clone RP11-353K12 map 8, WORKING DRAFT  
SEQUENCE, 25 unordered pieces.  
AC025853  
AC025853 3 GI:8077000  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 172162)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 8, clone RP11-353K12  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 172162)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgaier,B., Brown,A., Burkett,G.,  
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,R.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 25, 2000 this sequence version replaced gi:7342114.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L7454

Center Clone name: 353\_K\_12

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156061 bases at least Q40

Consensus quality: 163284 bases at least Q30

Consensus quality: 166861 bases at least Q20

Insert size: 182000; agarose-fp

Quality coverage: 4.0 in Q20 bases; agarose-fp

Quality coverage: 4.3 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1104: contig of 1104 bp in length  
\* 1105 1204: gap of 100 bp  
\* 1205 3226: contig of 2022 bp in length  
\* 3227 3326: gap of 100 bp  
\* 3327 4615: contig of 1289 bp in length  
\* 4616 4715: gap of 100 bp  
\* 4716 6570: contig of 1855 bp in length  
\* 6571 6670: gap of 100 bp  
\* 6671 8592: contig of 1922 bp in length  
\* 8593 8692: gap of 100 bp  
\* 8693 11929: contig of 3237 bp in length  
\* 11930 12029: gap of 100 bp  
\* 12030 15549: contig of 3520 bp in length  
\* 15550 15649: gap of 100 bp  
\* 15650 19338: contig of 3689 bp in length  
\* 19339 19438: gap of 100 bp  
\* 19439 23177: contig of 3739 bp in length  
\* 23178 23277: gap of 100 bp  
\* 23278 29032: contig of 5755 bp in length  
\* 29033 29132: gap of 100 bp  
\* 29133 34839: contig of 5707 bp in length  
\* 34840 34939: gap of 100 bp  
\* 34940 40597: contig of 5658 bp in length  
\* 40598 40697: gap of 100 bp  
\* 40698 47278: contig of 6591 bp in length  
\* 47279 47378: gap of 100 bp  
\* 47379 54796: contig of 7418 bp in length  
\* 54797 54896: gap of 100 bp  
\* 54897 59726: contig of 4830 bp in length  
\* 59727 59826: gap of 100 bp  
\* 59827 68926: contig of 9100 bp in length  
\* 68927 69026: gap of 100 bp  
\* 69027 75947: contig of 6921 bp in length  
\* 75948 76047: gap of 100 bp  
\* 76048 83423: contig of 7376 bp in length  
\* 83424 83523: gap of 100 bp

\* 83524 90514: contig of 6991 bp in length  
\* 90515 90614: gap of 100 bp  
\* 90615 99034: contig of 8420 bp in length  
\* 99035 99134: gap of 100 bp  
\* 99135 109649: contig of 10515 bp in length  
\* 109650 109749: gap of 100 bp  
\* 109750 122299: contig of 12550 bp in length  
\* 122300 122399: gap of 100 bp  
\* 122400 137369: contig of 14970 bp in length  
\* 137370 137469: gap of 100 bp  
\* 137470 154199: contig of 16730 bp in length  
\* 154200 154299: gap of 100 bp  
\* 154300 172162: contig of 17863 bp in length.  
FEATURES  
source  
1. 172162  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-353K12"  
/clone\_lib="RPCI-11 Human Male BAC"  
1. 1104  
/note="assembly\_fragment"  
misc\_feature  
1205..3226  
/note="assembly\_fragment"  
misc\_feature  
3327..4615  
/note="assembly\_fragment"  
misc\_feature  
4716..6570  
/note="assembly\_fragment"  
misc\_feature  
6671..8592  
/note="assembly\_fragment"  
misc\_feature  
8693..11929  
/note="assembly\_fragment"  
misc\_feature  
12030..15549  
/note="assembly\_fragment"  
misc\_feature  
15650..19338  
/note="assembly\_fragment"  
misc\_feature  
19439..23177  
/note="assembly\_fragment"  
clone\_end:N7  
vector\_side:right  
misc\_feature  
23278..29032  
/note="assembly\_fragment"  
misc\_feature  
29133..34839  
/note="assembly\_fragment"  
misc\_feature  
34940..40597  
/note="assembly\_fragment"  
misc\_feature  
40698..47278  
/note="assembly\_fragment"  
misc\_feature  
47379..54796  
/note="assembly\_fragment"  
misc\_feature  
54897..59726  
/note="assembly\_fragment"  
misc\_feature  
59827..68926  
/note="assembly\_fragment"  
misc\_feature  
69027..75947  
/note="assembly\_fragment"  
misc\_feature  
76048..83423  
/note="assembly\_fragment"  
misc\_feature  
83524..90514  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
misc\_feature  
90615..99034  
/note="assembly\_fragment"  
misc\_feature  
99135..109649  
/note="assembly\_fragment"  
misc\_feature  
109750..122299  
/note="assembly\_fragment"  
misc\_feature  
122400..137369  
/note="assembly\_fragment"  
misc\_feature  
137470..154199  
/note="assembly\_fragment"



misc_feature		154300..172162			
/note="assembly_fragment"					
BASE COUNT		47201 a 37955 c 38816 g 45788 t		2402 others	
ORIGIN					
Query Match 54.5%; Score 4926.8; DB 64; Length 172162;					
Best Local Similarity 99.1%; Pred. No. 0;					
Matches 4999; Conservative 3; Mismatches 20; Indels 22; Gaps 4;					
QY	4025	gcaatgggggtatagtgattaaactgttgggatcaaaatctctactcttatactttatattt	4084		
DB	99013	GCAATGGGGAATAG-GAATAAACTG-TGGGATCAAATCTACTCTTAACTTTTTTTT	98956		
QY	4085	tatatatatatatata-----atatatatatatattagccctca	4125		
DB	98955	TAGCCCTCA	98896		
QY	4126	ggctggctcacttcacagctgtttgctatcataacctctctgtcctcagtttcatgtat	4185		
DB	98895	GGCTGGTCACTTCACAGCAATTTGCTATCATAACTCTCTGTGCCCTCAGTTTCATTGAT	98836		
QY	4186	gtaattaggagactactaatagtaactcttcaatcggttgtaaggaatagatagacaaa	4245		
DB	98835	GTAAATTTGAGGACTACTAATAGTACTACTTTCATCGGGTTGTAAGGAATAGATGAGCAA	98776		
QY	4246	tgtatggcttggcacttaataacactcaaaattattagtaaaagtattgttataataa	4305		
DB	98775	TGTATGGCTTGGCACTTAATAACACTACAAATATTAGTGAAGATGTTTATATAATA	98716		
QY	4306	tacttctgttggtgagcgtgtggtgcacgcctgcaatcccgagcaactttgggagcag	4365		
DB	98715	TACTTCTGTGTGGCTAGGCGTGGTGCCTCACGCCCTGCAATCCACGACCTTTGGAGGCAG	98656		
QY	4366	agggcaggcagcaacttgaggtcaggaattcgagatcagcctggtcccaacatgaggaacc	4425		
DB	98655	AGGCAGGCAGAGCACTTGAGGTTCAGGAATTCGAGATCGACCTTGGCCAAACATGAGGAACC	98596		
QY	4426	ccgtctctactaaaaatacaaaaatcagccagcagatggtggcaggtgtctgtaatccag	4485		
DB	98595	CCGTCTCTACTAAAAATAACAAAATCAGCCAGGCATGTTGGCAGGTGCTGTAAATCCCGAG	98536		
QY	4486	ctactttggaggctgaggcaggagaatcagagggaggcggagggttcagtgagccaaga	4545		
DB	98535	CTACTTGGGAGGCTGAGGCAGAGGAATCAGAGGGGAGCGGAGGTTCAGTGCAGCCNAGA	98476		
QY	4546	tcacgccactacacccccagcctagtgtaaaaagcagacctctcaaatattacaataat	4605		
DB	98475	TCACGCCACTACACCCAGCCTTAGGTGACAAAAGCGAGACTTCTCAAAATAATAACAATAAT	98416		
QY	4606	aataactatgttcattatatacatgatattatttttactatttactatagccta	4665		
DB	98415	AATATACTATGTGTCATTATACATCATTAATTAATTAATTAATTAATTAATTAATTAATTA	98356		
QY	4666	gtcgtataacctgggaaaaaggtcacagaatgttcagcttacttcagattggacaaaag	4725		
DB	98355	GCTCGAATACCTTGGGAAAAGGTTCACGCAATGTTCAGCTTACTTTCAGATTGGACAAAG	98296		
QY	4726	gtggaatgcctaaacacggggccacccgatccggagtggttatttttaggcagctg	4785		
DB	98295	GCTGGAATGCCCTAAACACCGGGCCACCGCATCCGGAGTGGCTTGGTTATTTTAGGCAGCTG	98236		
QY	4786	agctgtcacttcccctgggttaaggacaactcaactctgtggcactctgtctccacccccacct	4845		
DB	98235	AGCTGTCACTTCCCTTGGGTAAAGGACACTCACCTCTTGGCACTCTGTCTCCACCCACCCCT	98176		
QY	4846	cggcaggtgtgcagaagtccaggcgagatctccctcctgaaagcagcagctgaaggagtc	4905		
DB	98175	CGCAGGTGTGCCAAGTTCAGCGGAGATCTCCCTCCTGAAGCAGCAGCTGAAGGAGTCC	98116		
QY	4906	cagacggaggtgaacgcgaaggttagcgagatcctgggtctcaaggcacagctgaaggac	4965		
DB	98115	CAGACGGAGGTGAACGCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACAGCTGAAGGAC	98056		

QY	4966	acgcggggcaagctgagggcctgagctgagctgagacccaggacctgagggcgccctgcgc	5025	
DB	98055	ACGCGGGGCAAGCTGAGGGCCTGGAGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGC	97996	
QY	5026	accaagggcctggagctgaggtctgtgagaatgagctgcagcgcaagaagcaacagggcg	5085	
DB	97995	ACCAAGGGCCTGGAGCTGAGGTCTGTGAGAAATGAGTGCAGCGCAAGAAGAACGAGGCG	97936	
QY	5086	gagctgtcgcgggagaaggtgaaacctgtggagaaggagctgcagagctcggggcccag	5145	
DB	97935	GAGCTGCTGCGGGAGAAGGTGAACCTGTGGAGGAGGAGCTGCAGAGCTGCGGGCCCCAG	97876	
QY	5146	gcgcgccttggcccgcgacatggggcgccaccttccccgagagcgtccctcctcctgacg	5205	
DB	97875	GCGGCCCTTGGCCCGGACATGGGGCGCCACCCTTCCCGAGGAGCTCCCTGCCCTGCAG	97816	
QY	5206	cggggagctggagcggctcgggcccgagctcgggggaggagcgcaaggcccatgaccagatg	5265	
DB	97815	CGGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGCAAGGCCATCACCAGATG	97756	
QY	5266	tactcgggcttcacagctagcggctcgtgtggaaggaggagaagaagtgattcaag	5325	
DB	97755	TCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAAGAGGTGATTCA	97696	
QY	5326	taccagaacaagctcagcagagctacgtggccatgtaccagcggaaccagcgctggag	5385	
DB	97695	TACCAGAACAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGCGGAACACGCGCCCTGGAG	97636	
QY	5386	aaggcccttcgagcagctggcagcgtgggacagcgccggggagcccttggaggttgacctg	5445	
DB	97635	AAGGCCCTTGCACAGCTGGCAGCTGGGACAGCGCGGGGAGCCCTTGGAGGTTGACCTG	97576	
QY	5446	gaaggggctgacatccccacagaggacatcatagcactgagatctgaggggctgcctgg	5505	
DB	97575	GAAGGGGCTGACATCCCTACGAGGACATCATAGGCACCTGAGATTGTAGGGGCTGCCTGG	97516	
QY	5506	gaagcgagctcgggacctggcactggagcgagggctctcccgatgcatccccctgct	5565	
DB	97515	GAAGCGGAGCTGTGGGACCTGSCACTGGGAGCGAGGGCTCTCCCGTGCAATCCCCCTGGT	97456	
QY	5566	cagcaattcagaccctctgagagcgcactccctgggacacagaccagagaccgccga	5625	
DB	97455	CAGCAATTTCAGACCCCTTTGAGAGACGCCACTCCCTGGGACACAGACCAGACCCCGA	97396	
QY	5626	ggggggcaggatggccttcttccctctctgtatgtcccaagtctcaaccagccctga	5685	
DB	97395	GGGGAGGCGAGGATGGCCTTCTCTCCCTCTCTGATGTGCCAGTGTCTACACGACCTGCA	97336	
QY	5686	gcccaaccagagctcagccctgaactcctctgcttccagagatgggtccagggctct	5745	
DB	97335	GCCCAACAGACCTGAGGCCCTGACTCTCTGGCTTTTCCAGGAGATGGGTCCAGGGGTCT	97276	
QY	5746	gtcgtcttggtaaaggcctccctaaaactttggcctttgttcgaaatagatatcctctcc	5805	
DB	97275	GTCTGCTTGGTTAAGGCTCCCTAAACTTTTGGCCTTTGTTCGAAATAGATATCTCTCC	97216	
QY	5806	ccctcctccagggaagggtggccacagcaagaacagcggtccctccctcctctctcatcca	5865	
DB	97215	CCCTCTCTCCAGGGAAGGTGGCCACAGCAAGTACAGCGGCTCCCTCTGTCTTCTCATCCCA	97156	
QY	5866	accttttttctcctgcacacatiggaaatgccttggaaatagaaagaagccatatatga	5925	
DB	97155	ACCTCTTTTCTCTCTGGACATTTGGAAATGCTTGGAAATAGAAAGAACCATATATGA	97096	
QY	5926	ccagaagccttggaaaccagcccatcagaacctgagctatttctcctggccgcgagagt	5985	
DB	97095	CCAGAAGCCTTGGAAACCAAGCCCATCAGAACTGAGCTATTTTCTCTGCGCCGACAGGT	97036	
QY	5986	gtaggggtggaatgagccggggggaagctggtttgaaacctcaggggtgtccccagcccc	6045	
DB	97035	GTAGGGGTGAATGATGCGCGGGGAAGCTGGCTTTTGAACCTCAGGGCTGTCTCCAGCCCC	96976	







Db	1671	GATGTCTCTCGGGTTCACAGATGAGCGGCTCTGTGTGAAGAGGAGGAGAAAGGTGAT	1730
Qy	5322	tcagttaccagaataacagctgcagcagagctacgtgqccatctaccagcggaaccagcgccct	5381
Db	1731	TCACTACCAGAAACAGCTGCACGACAGGTACGTGGCCATGTACCAGCGGAACACGCGCCT	1790
Qy	5382	ggagaagggccctcagcagctggcacgtgggcacgtgggacagcccgggagcccttggaggttga	5441
Db	1791	GGAGAAGGCCCTCAGACAGCTGGCACAGTGGGACAGCCGCCGGAGAGCCCTTGGAGGTGA	1850
Qy	5442	ccctggaagggctgacatccctcagagagacatcagccactagatcatcagagggctgc	5501
Db	1851	CTGTGAAGGGCTGACATCCCTACAGAGACATATACCCACTGAGATCTGAGGGGTGC	1910
Qy	5502	ctgggaagggcagctctggggaacctggcactggagcagggctctccctgcatccccccc	5561
Db	1911	CTGGGAAGGGCAGCTCTGGGACCTGSCACTGGGAGGCGAGGCTCTCCGTGTCATCCCCC	1970
Qy	5562	tgtcagcaaatcagacccctctgagagacgcacatccctgggacacagaccagagacc	5621
Db	1971	TGCTACAGCAATTACAGACCCCTCTGAGAGACGCACTCCCTGGGACACAGACCCAGGACC	2030
Qy	5622	ccaggggagggcagatgagcttctccctctctcatccatgctcccatgctccacagccc	5681
Db	2031	CCGAGGGAGGGCAGATGGCCCTTCTTCCCTCTCTGATGTCCAGTGCTACACGCC	2090
Qy	5682	tgcagccacacagctcagggccctgaactctctggcttcccaggagatgggtccaggg	5741
Db	2091	TGCAGGCCACACAGACGTCAGGCCCTGACTCTCTGGCTTTCCAGGAGATGGGTCCAGGG	2150
Qy	5742	gtctgtctgttggttaagggctccctaaactttggcccttgttcgaaatagatatccct	5801
Db	2151	GTCTGTCTGTCTTGGTTAAGGGCTCCCTAAACTTTTGGGCTTTGTGTGAAATAGATACTCT	2210
Qy	5802	ctccctctctccaggaaggtggccacagacaagacagcgctccctcgcttctcat	5861
Db	2211	CTCCCCCTCTCCAGGGAGGTTGGCCACAGCAAGAACAGCGGCTCCCCCTCGCTTCTCAT	2270
Qy	5862	cccaaccttttctctctggacaccttggaaatgccttggaaatagaaagaagccatat	5921
Db	2271	CCCAACCTCTTTTCTCTCTGGACACATTGGAATGCCTTTGAAATAGAAAGAGCCATAT	2330
Qy	5922	atgacaagaaccttggaaaccagcccatcagaacctgagctatttctctggccgcag	5981
Db	2331	ATGACCAGAAGCCTTGGACACAGCCCACTACAGACCTCAGACTATTTTCTCTGGCCGCAG	2390
Qy	5982	aggttaggggtggaatgagccgcgggaagctggcttgaacctcagggctgtcccg	6041
Db	2391	AGGTGTAGGGGTGGATGAGCCCGGGGAAGCTGGCTTTGAAACCTCAGGGCTGTGCCAG	2450
Qy	6042	ccccgcgaagccaggaaggggagagacagcagccagctgagtgtagaacctctgc	6101
Db	2451	CCCCGGCAAGCCACAGGAAGGGGGAGAGACAGCACGCCACAGAGTGTGGAGACCTGCG	2510
Qy	6102	cacagccagaggggcagaggagaaatccaaaggggttgagagccagtggcgggttgatggc	6161
Db	2511	CACAGCCAGAGGGCAGAGGGAGAAATCCAAGGGTTGAGAGCCAGTGCGGGGTGATGGC	2570
Qy	6162	cagcccttggggcccaagccctgtttactgttcttgcaaatggagagctgagcagcctct	6221
Db	2571	CAGCCCCCTGGGGCCCAAGCCCCCTGTTTACTGTGTTCTTGAAATGGGAGCTCAGCAGCCTCT	2630
Qy	6222	ggacagccagtgcaccttgcactcggtagccactctcttaagccataaacctcaggc	6281
Db	2631	GGACAGCCAGTGACCTTTGACCTCGGTGACCACTCTCTTTAAGCCATAGACCCCTGAGGC	2690
Qy	6282	cctgggctgggtgctgggaaggaggggttgaaccacacgtagaacacagaggggtgtagcttt	6341
Db	2691	CTTGGGCTGGGTGCTGGGAAGGAGGGTTGAAACCAACCGTCAACAGAGGGGTGTGGCTTT	2750
Qy	6342	ccagkacacctcaggagcctcccatctgttcacgtctggggccagaggtctggagtcctct	6401
Db	2751	CACGAGACCCCTCAGGAGAGCCTCCCATCTGTCCAGCTGGGCCAGAGGCTGGGAGTGCCTT	2810

Qy	6402	acctgcttcacgttggccggcggtcactctggaatgtttttccctccccagaatacaagct	6461
Db	2811	ACCTGCTTACGTTGGCGCGCGCTACTCTGGAATGTTTTTCCCTCCCCAGAATCAAGCT	2870
Qy	6462	tttcttgatccagaagagcccatatcaactaagatggcaataatgtatctgggcatttt	6521
Db	2871	TTTCTCTTGATGCCAAGAGAGCCCATATCACTAAGATGGCATATATGTGATCTGGGCATTTT	2930
Qy	6522	cctcctcgctacagccaggttagcgcaaacctttcccccctttagcaaccttcagggct	6581
Db	2931	CCTCCTCTGCTACAGCCAGGTTTACGGGCAAAACCTTCCCCCTTAGCACCTTCAGGGCT	2990
Qy	6582	gagttctgggtttctagaggtcaggacggtcctcagagcgccaggaagccagagcccca	6641
Db	2991	GAGTTCTGGGTTTCTAGAGTCAAGGACGGCTCTCTCAGAGCGCCAGGAAGCCAGGCCCA	3050
Qy	6642	agaagcagcaaaagagagatacacacagcagtgtaatatagactggccacacagccatcct	6701
Db	3051	AGCAGGACGAAAAGAGGCATACACACAGCAGTGTGTAATAGCTCTGGCCACCCAGCCATCCT	3110
Qy	6702	ccctccacctcaagacccccattgtccsagactaaagatccagagagcagctcccttt	6761
Db	3111	CCCTCCACCTCAGACCCCATTTGTCCACAGCTAAAGGATCCAGAGAGAGCTCCCTTT	3170
Qy	6762	ctcaggagctgggcagtgccccagggagtccagggtttctctgcagatgtcgcgagcgg	6821
Db	3171	CTCAGGAGCTTGGCAGTGCCCCAGGGAGTTCAGGGTTTCTCTGCAGATGTGCGAGCGG	3230
Qy	6822	gagcggctggttagagagagataaaagtgtaggtttctctgttctgttctgcaggatttt	6881
Db	3231	GAGCGGTGTGTAGAGAGAGATAAAAGGTGGAGTTCTCTGTGTTTGGTTTCAGGGATTTT	3290
Qy	6882	attttaatttatgacagcaggtctgtctgtcccccagcgtagagtgcaagtgcagtcg	6941
Db	3291	ATTTTAAATTTATGACAGAGGCTTGTCTGTCTCCCCAGGCTGGAGTGCAGTGGCATG	3350
Qy	6942	atcatgctcactgcagctcatctactcctgggctcaagcaatccctcctgcctcagccttc	7001
Db	3351	ATCATAGTCACTGCAGCCCTACATCCTTGGGCTCAAGCAATCCTCCTGCCTCAGCCCTC	3410
Qy	7002	caactagctgggactacaagtcgcgcgcacccgtgcctgcgtcaacttttcatttttgt	7061
Db	3411	CAACTAGCTGGGACTACAGTGTGGCGGCCACCGTCCGTGCGCTAACTTTTCAATTTTTTGT	3470
Qy	7062	aggacggggtctcgtttgttgccaaaagctagtcctcaaatgtggcctcaagcaatcc	7121
Db	3471	AGGACGGGGTCTCGTTTTGTGCCAAAGCTGGTCTCAACCTGTGGGCTCAAGCAATCC	3530
Qy	7122	acctgccttggcctcccaagtgctgagattgcagatgtgagccaccgtgcctggccaga	7181
Db	3531	ACCTGCCTTGGCCTCCCAAAGTCTCAGATTGCAGATGTGAGCCACCGTGCCTGGCCAGA	3590
Qy	7182	ttttcttttattcttctttttctttttttgttcttctgtcttcttcttcttcttcttctt	7241
Db	3591	TTTTTCTTTTATCTCTTTTCTTTTTTGTCTTTTGTCTTTTCTTTTCTTTTCTTTTCTTTT	3650
Qy	7242	agacctagagcgtgttccatgttctatttttgactgttagccaaagctgcgttctcagg	7301
Db	3651	AGACCTAGCAGGCTGTTCATGTTCTATTTTTTACTGTAGCCACACAGCTGCTGTCTCAGG	3710
Qy	7302	acagcatcccttccacatgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgc	7361
Db	3711	ACACATCCCTTCCACATGCCCTGCCCTGCTGCGCTGCTGAGATGAGAGGGGAGCGTCT	3770
Qy	7362	gggaacctgcagttccaaagccagtcctcccatttctgcctcgtctcaccgctggccctaga	7421
Db	3771	GGGAACCTTGCGAGTCCAAGGCCAGTCCCCATTTCTGCCTCGCTCACCGCTGGCCCTTAGA	3830
Qy	7422	gaccccgaggtaggggtgggaagatgctctctcttcttcccccccgccctcatgggtccta	7481
Db	3831	GACCCCGAGGTAGGGGTGGGAGATGCTTCTCTCTTGGCCCCCGCCCTCATGGGTGCTTA	3890

QY 7482 gcccttccctgagtgagggtgagccagagtcacaccttttctgtgctggtctaccttc 7541  
Db 3891 GCCCTTCCCTGAGTGGGGCTGAGGCCAGAGTCACCTTTCTGTGCTGCTTACCTTC 3950  
QY 7542 ctgtccctgaggttaaacggtgcccatctctgcatctctaaacacagagagctttct 7601  
Db 3951 CTGTCCCTGAGGTTAAAGGGTGCCATCTCTGCCATCTCTAAACGACGAGAGAGCTTTCT 4010  
QY 7602 ggaatttcaaacattgtcttagtcccaagctaggcttaaacctggaaattctacaagcca 7661  
Db 4011 GGAATTTCAACCAATGCTCTTAGTCCCAAGCTAGCTTAAACCTGGAACTTACAAGCCA 4070  
QY 7662 aaagtccctccctgctgagggcagtcacctccattggtgcagtcacagaccagaagtcac 7721  
Db 4071 AAAGTCCCTCCCTGCTGAGGGCAGTACCCTCCATTGGGCACAGTCCAGACCCCAAGTCAA 4130  
QY 7722 agatgccccattccttgctgctcagctcagctcagctcagctcagctcagctcagctc 7781  
Db 4131 AGATGCCCAATTCCTTGGGCCCTCAGGCCCTCAGTTCTCTTCAATTCACCAAGGCCGTGCCCT 4190  
QY 7782 gtttagtttttccctccagtgagactgccccccagcagagacagagaaaggctgctccc 7841  
Db 4191 GTTTAGTTTTTCTCCAGTGAGACTGCCCCCAGGAGACAGAGAAAGGGCTGGCTCCC 4250  
QY 7842 cctcccaggtggagaccccccccaactccaggaaagagcagtcagagtcagtgctct 7901  
Db 4251 CTTCCCGAGGCTGGAGACCCCCCAACTCCAGGAAGAGCAGTCAGAGTCCAGTGCTCT 4310  
QY 7902 gctcagacgttgctcagaaagaagtgtcgtccacacccagggggaagccctgagggcga 7961  
Db 4311 GCCTCAGAGCTTGCTGAGAGAAAGTGGCTGCCACACCCAGGGGAAGGCCCTGAGGCGGA 4370  
QY 7962 ggtctgtctccagctggtctccagtgaccttcctacacagagagtgacagctctctcc 8021  
Db 4371 GGTCTGTCTCCGCCATGGTGTCCCCGGTACCTTCCATACACAGGAGTGCAGCCCTTCTCC 4430  
QY 8022 atatctcattggcctgtcccagggcgccagatgtgtccccccagggcctgtgctctac 8081  
Db 4431 ATATCTCATGCCCCCTGTCCAGGCGCGCCAGATGTGTCCCCCCAGGCCCTGTGCTTAC 4490  
QY 8082 gtccaaagtgagatgtcttccctgggctgccacacagcccccccgccagagtgcccaac 8141  
Db 4491 GTCCAAAGTGGCAGATGTCTTCCCTGGGCTGCCACACGCCGCCGCCAGAGTGGCCAC 4550  
QY 8142 cgtggcactagaatcaagatcctcgtgacettgcaactcactctcctgtggtgtct 8201  
Db 4551 CGTGGCACTAGAATCAAGTAGTCTTCGCACTTGCACCTACCTTCTGTGGGTGTCT 4610  
QY 8202 ttctgcccctgtccaaaagcgccctcactattcttggaccatgccagattctgcctctct 8261  
Db 4611 TTCTTGCCCTGTCCAAAAGCGCCCTCACTATTCTTTGGACCATGCCAGATTCTGCCTCTCT 4670  
QY 8262 gaaagagagctctgacagcagaagcctccaaagcagacagcctgtgcccagggccccagac 8321  
Db 4671 GGAAGAAGGCTCTGGACAGCAGAGAGCCCTCCAAAGCAGACAGAGCCTGGCCCCCAGAC 4730  
QY 8322 aggtgggcttccctcctcctcctggtgagcctgctggtgagcctggtgagcctcagcactc 8381  
Db 4731 AGGGTGGGCTTCTGCCCTTCCCTCTGGGACGCTGTGTGGCCGACCCACTTACCCACTC 4790  
QY 8382 ggaaggacaaactcctctgtgtcccaagagcgctcgtcagggagagcagcactccgcat 8441  
Db 4791 GGATGGACCAACTGCTGTGTCCCAAGGACGCTGTGAGGAGAGAGCAGCACTCCGCAT 4850  
QY 8442 caccctcaacaaggaatcgactctgcccctggacctggaacctggaaagactggactgtcaggggt 8501  
Db 4851 CACCTCACCAAGGATCGGACTCTGCCCTTGGACCTTGGAAACGACTTGGACTGTACGGGGT 4910  
QY 8502 tccctcctagctctccagtgactcctcctcagcgacacacagccccctatagcactgac 8561  
Db 4911 TCCCTCTAGCTCTCCAGTAGTCTCTTCCAGGCACTCTTCCAGGACACACAGCCCCCTATAGCATGAC 4970  
QY 8562 tcacatgggactgggatatggggggcatctcttccccccagagagggcactcagtgagcctct 8621

Db 4971 TCACATGGGACTGGGATATGGGGGCATCTCTTCCCAGAGAGGCACTCAGTCAGGCTCCT 5030  
QY 8622 gtgcctgccccagctctgggccaatctcttagtgagacagtgccgaaactaaagccagg 8681  
Db 5031 GTGCTGCCCCCAGCTGTGGGCATCTCTTAGGTGAGACAGTTGCCGAAACTAAAGCCAGG 5090  
QY 8682 cctgctgagagcagcagcttgaggagaggttccctcgacacctcaagccatcat 8741  
Db 5091 CTTGCTGGAGGAGCAGCAGCTTGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCAT 5150  
QY 8742 ggggtggtgtgctgcatgacagagctgcacccctgggccaagcggggtgctcaccacc 8801  
Db 5151 GGGGTGGGTGCTGCCATGACAGAGCTGCACCCCTGGGCCAGCGGGCTGCTCACCCACC 5210  
QY 8802 tcttgcaagtggtccttctgtgcgctcagcagagctggagccccccagagagg 8861  
Db 5211 TCTTGTGAAGTGGCCTTTGTGCTGCGCTCAGGAGCTGGAGCCCGCCAGCAGAGG 5270  
QY 8862 caggctggagcagcagcagctctggaagatgtacatagttattttctcttctgtggttc 8921  
Db 5271 CAGGCTGGAGCGACAGCATCTGGAAGATGTACATAGTATTTTCTCTTGTGGTTTC 5330  
QY 8922 ttgttgggtgtgtgcttttgacagcttcttatttttgacgtcactttttggcca 8981  
Db 5331 TTGTTTGGTTGCTTGTGCTTTTGACAGCTTCTATTTTGTGACGTCACTTTTGGCCA 5390  
QY 8982 tgaactatttggtggcaattttatgtttttatttatgaataaagaatgccattctcac 9041  
Db 5391 TGTAACATATTGTGCAATTTTATGTTTATGATGAATAAAGAATGCCATTTCTCAC 5450  
QY 9042 gccctct 9048  
Db 5451 GCCCTCT 5457

RESULT 4  
AF123656 1515 bp mRNA PRI 07-APR-1999  
LOCUS Homo sapiens clone D14 FEZ1 (FEZ1) mRNA, alternatively spliced,  
DEFINITION complete cds.  
ACCESSION AF123656  
VERSION AF123656.1 GI:4572469  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1515)  
AUTHORS Mori,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Ishii,H., Baffa,R., Alder,H. and Croce,C.M.  
The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,  
and its expression is altered in multiple human tumors  
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
99199287  
REFERENCE 2 (bases 1 to 1515)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
Direct Submision  
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
JOURNAL Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
source 1..1515  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/clone="D14"  
1..1515  
/gene="FEZ1"  
1..1515  
/gene="FEZ1"  
/codon\_start=1

```
/product="FEZ1"
/protein_id="AAD23837.1"
/db_xref="GI:4572470"
/translation="MGSVSLISGSHFHKCRASQYKLRKSHLKLKLNRYSDGLLRF
GFSQDSGHGKSSKMGKSEDFYIKVSOAKRGSHHPDYATLSSGDLGGAGVDFDPST
PPKLMPPFNQLEMGSEKAVRTAFKPVLPKRSAILHSSPESASQHLHPAPDPKPEQ
ELKPGICGALSDFGRNMSLSLPTSSSYQLDPLVTPVPTSRFGSAHNITQGIY
LQDSNMSLKALSFSDGSKLGHKNADKGPSCVSPISITDECSTQELQKLLERGA
LQKLSRPEEKELASLAVEERPRRCDELEGPEPKGNKQKQASQKSORAQVHLQ
VLOLOQEKROLROELESIMKEODLLETKLRSYERKTSFGPALETOEWVCQKSGEIS
LLKQOLKESQTEVNAKASEILGLKQKLDTRGKLSGLRLTODLEGALRTKGLLEVC
ENELQRKNEAEALLREKHLVWKEEKVIQYQKQLOQSYVAMYQVORNRLEKALQOL
ARGDSAGEPLEVDLEGADIPYDIATET"
BASE COUNT 336 a 483 c 482 g 214 t
ORIGIN
Query Match 8.9%; Score 807.2; DB 10; Length 1515;
Best Local Similarity 99.6%; Pred. No. 2.6e-119;
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1646 gggctccgagaagggtgcagtgagccacagcttcaagcctgtgtcgcacggtcagg 1705
Db 345 GGGCTCGGAGAGGGTGCAGTGAGGCCACACAGCTTCAAGCCTGTGCTGCCACGGTCAGG 404
QY 1706 agccatctgcactctccccgagagtgccagccaccagctgcacccgcgcctccaga 1765
Db 405 AGGCATCTGCACCTCTCCCGGAGAGTGCCAGCCACCAGCTGCACCCGCCCTCCAGA 464
QY 1766 caagcccaaggagcagctgaagcctggcctgtgtctgtgggcgtgtcagactccgg 1825
Db 465 CAAGCCCAAGGAGCAGAGCTGAAGCCTGGCCTGTCTGTGGGGCGCTGTCAAGACTCCGG 524
QY 1826 ccggaactcatcttcagctgccccacacacagcagcagcagctgcacccgcgcctccaga 1765
Db 405 AGGCATCTGCACCTCTCCCGGAGAGTGCCAGCCACCAGCTGCACCCGCCCTCCAGA 464
QY 1766 caagcccaaggagcagctgaagcctggcctgtgtctgtgggcgtgtcagactccgg 1825
Db 465 CAAGCCCAAGGAGCAGAGCTGAAGCCTGGCCTGTCTGTGGGGCGCTGTCAAGACTCCGG 524
QY 1886 gctgggtcacacccctgggaccacaagccgttttggggctcgcgcacacaacacaccca 1945
Db 585 GCTGGTCCACACCCGTGGGACCCACAAGCCGTTTGGGGCTCGCCGCCACAACATCACCCA 644
QY 1946 gggcatcgtctccagagacagaacaatgatagcctgaaggctctgtctcttctccagcgg 2005
Db 645 GGGCATCTGCTCTCCAGGACAGCAACATGATGAGCCCTGAAGGCTCTGTCTCTTCCGACGG 704
QY 2006 aggtagcaagctgggcactcaacaagcagcagcagcagcagcctgtgtcgtcgcctccc 2065
Db 705 AGGTACAAGCTGGGCCACTCGAACAAGCAGACAAGGCCCTCGTGTGTCCGCTCCCC 764
QY 2066 catctccagcagcagtgagcagcatccagagctgggaacagaagctgttgagagggaggg 2125
Db 765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGAAGCTGTGGAGAGGGAGG 824
QY 2126 cgcctccagaagctcagcagcagctttgaggagaaggagcttgctccagcctggccta 2185
Db 825 CGCCCTCCAGAAGCTCGACGCGCAGCTTGAAGGAGAGGAGCTTGCCTCCAGCCTGCCTA 884
QY 2186 cgagcgagcgcgcgcgtcgaggaacagctggagggcgcgcgcgcgcgcgcgcgcgcgc 2245
Db 885 CGAGGAGGCCGCCGCGCTGCAGGACAGCAGCTGGAGGGCCCGGAGCCCAAGAGCCGCA 944
QY 2246 caagctcaagcagcgcctcgcagagaagcagcgcgcgcgcgcgcgcgcgcgcgcgcgc 2305
Db 945 CAAGCTCAAGCAGGCTCGCAGAAGAGCCAGCGCGCGCAGCAGGCTCTGCACCTGCAGGT 1004
QY 2306 actgcagcttcagcagagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2365
Db 1005 ACTGCAGCTTCAGCAGGAGAGCGCAGCTCCGGCAGGAGCTCGAGAGGCTCATGAAGA 1064
QY 2366 gcagagacctgctgagacaagctcaggttcctacgagagggagagacacacagcttcggccc 2425
Db 1065 GCAGGACCTGCTGGAGACCAAGCTCAGGTCCCTACGAGAGGGAGAGACCAAGCTTCGGCCC 1124
QY 2426 cgcgtggagagagacccagtgaggaggtgaggc 2457
```

```
|||||
Db 1125 CGCGCTGGAGGACACCCAGTGGGAGGTGTGCC 1156
RESULT 5
AF123655 1614 bp mRNA PRI 07-APR-1999
Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, alternatively
spliced, complete cds.
AF123655
AF123655.1 GI:4572467
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1614)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
REFERENCE 2 (bases 1 to 1614)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 2335 10th street, Philadelphia, PA 19107, USA
FEATURES
Location/Qualifiers
1..1614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="T8D145M4"
1..1614
/gene="FEZ1"
1..1614
/gene="FEZ1"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23836.1"
/db_xref="GI:4572468"
/translation="MGSVSLISGSHFHKCRASQYKLRKSHLKLKLNRYSDGLLRF
GFSQDSGHGKSSKMGKSEDFYIKVSOAKRGSHHPDYATLSSGDLGGAGVDFDPST
PPKLMPPFNQLEMGSEKAVRTAFKPVLPKRSAILHSSPESASQHLHPAPDPKPEQ
ELKPGICGALSDFGRNMSLSLPTSSSYQLDPLVTPVPTSRFGSAHNITQGIY
LQDSNMSLKALSFSDGSKLGHKNADKGPSCVSPISITDECSTQELQKLLERGA
LQKLSRPEEKELASLAVEERPRRCDELEGPEPKGNKQKQASQKSORAQVHLQ
VLOLOQEKROLROELESIMKEODLLETKLRSYERKTSFGPALETOEWVCQKSGEIS
LLKQOLKESQTEVNAKASEILGLKQKLDTRGKLSGLRLTODLEGALRTKGLLEVC
ENELQRKNEAEALLREKHLVWKEEKVIQYQKQLOQSYVAMYQVORNRLEKALQOL
ARGDSAGEPLEVDLEGADIPYDIATET"
BASE COUNT 367 a 500 c 522 g 225 t
ORIGIN
Query Match 8.9%; Score 807.2; DB 10; Length 1614;
Best Local Similarity 99.6%; Pred. No. 2.5e-119;
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1646 gggctccgagaagggtgcagtgagccacagcttcaagcctgtgtcgcacggtcagg 1705
Db 345 GGGCTCGGAGAGGGTGCAGTGAGGCCACACAGCTTCAAGCCTGTGCTGCCACGGTCAGG 404
QY 1706 agccatctgcactctccccgagagtgccagccaccagctgcacccgcgcctccaga 1765
Db 405 AGGCATCTGCACCTCTCCCGGAGAGTGCCAGCCACCAGCTGCACCCGCCCTCCAGA 464
QY 1766 caagcccaaggagcagcagctgaagcctggcctgtgtctgtgggcgtgtcagactccgg 1825
Db 465 CAAGCCCAAGGAGCAGAGCTGAAGCCTGGCCTGTCTGTGGGGCGCTGTCAAGACTCCGG 524
```

QY	1826	ccggaactccatgtccagctgcccacacacagcagcagcagctaccagctgaccc	1885
Db	525	CCGGAACCTCCATGTCCAGCTGCCACACACAGCACAGCAGCAGCTACAGCTGGACCC	584
QY	1886	gtgtgtcacaccctgtggaccacaaagccgttttgggggtccgccacacacaccca	1945
Db	585	GCTGTGTACACCCGTTGGGACCCACAAAGCCGTTTGGGGGCTCCGCCACAAACATCACCA	644
QY	1946	gggcatcgtctccaggacagacaaatgatgagctgaagctctgtctcttcagcagg	2005
Db	645	GGGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCCGACGG	704
QY	2006	aggtagaagctggggccactgaaacaaagcagacaagggccctcgtgtgtcgtccccc	2065
Db	705	AGGTAGCAAGCTTGGGCCACTCGAACAAGCAGACAAGGGCCCTCTGTGTCTCCGTCCC	764
QY	2066	catctccacggcagtgagcagcatccagagctggaacagaagctgttgagaggaggg	2125
Db	765	CATCTCCAGGACGAGTGCGAGCATCCAGGAGCTGGAGCAGAGAGCTGTTGAGAGGGAGGG	824
QY	2126	cgcctccagaagctgcagcgcagctttgagagaagagcttgctccagcctggccta	2185
Db	825	CGCCCTCCAGAAGCTGCAGCGAGCTTTTGGAGAGAAGAGCTTGGCTCCAGCGTGGCCTA	884
QY	2186	cgaggagcggcggcgtgaggaagcagctggaggcccgagagcccaagcgga	2245
Db	885	CGAGGAGCGCGCGCGCTGCAGGGACGAGCTGGAGGCGCCGGAGCCCAAGGGCGCA	944
QY	2246	caagctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2305
Db	945	CAAGCTCAAGCAGGCTCGCAGAAGACCCAGCGCGCAGCAGGCTCTGACCTGCAGT	1004
QY	2306	actgagcttcagcaggaagcagcagctccgagagagctcagagcctaatgaaga	2365
Db	1005	ACTGACGCTTCAGCAGGAAAGCGGAGCTCCGGCAGGAGCTCGAGAGGCTCATGAAGA	1064
QY	2366	gcaggacctgtggagacaaagctcaggtctctacagagaggagagaccagcttcggccc	2425
Db	1065	GCAGGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGGGAGAGACCAAGCTTCGGCCC	1124
QY	2426	cgcgtggaggagaccagctggaggtgagc	2457
Db	1125	CGCGTGGAGGAGACCCAGTGGGAGGTGTGCC	1156
RESULT	6		
AF123657	1692 bp	mRNA	PRI
LOCUS	Homo sapiens	clone G3611	FEZ1 (FEZ1) mRNA, alternatively spliced, complete cds.
DEFINITION	AF123657	GI:4572471	
ACCESSION	AF123657		
VERSION	AF123657.1		
KEYWORDS	human		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1692)		
AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.		
TITLE	The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)		
MEDLINE	99199287		
REFERENCE	2 (bases 1 to 1692)		
AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA		
FEATURES	Location/Qualifiers		
	1..1692		

/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="8"		
/map="8p22"		
/clone="G3611"		
gene	1..1692	
/gene="FEZ1"		
CDS	1..1692	
/codon_start=1		
/product="FEZ1"		
/protein_id="AAD23838.1"		
/db_xref="GI:4572472"		
/translation="MGSVSSLLISGHSFKHCRASQYKLRKSSHLKLNRYSDGLLRF		
GFSODSHGKSSKMGKSEDFYIKYQKARGSHHPDYTALSSDGLGQAGVDFDPST		
PPKLMPPFSNLEMGSEKAVPTAFKPVLRSGAILHSSPESASHOLHPAPDRKPEQ		
ELKPLCSGALSDDGRNSMSLPTSTSSYOLDPLVTPVPTSFSGSAINITLREGA		
LQDNMSLKALSFSDGSKLGHNKADKPSKVRSPITDECSIQEIQEQRKLLREGA		
LQKLQSPERELASSLAYERPRRCDELEGPPEKGNKLPKQASOKSQRQAQVHLQ		
VLQIQEKRLQRLQELSLMKQDLETLKLSYREKTSFGPALRETOWEQCKSGEIS		
LLKQQLKESOTEVNAKSEILGLKQAKDTGKLGLELRTQDLEGALRTGLELEVC		
ENELQRKNAEALLREKYNLLERLRAELREERQHDQSSGFQHERLWVKEKEKVIO		
YOKLQQSIVAMYQNRQLEKALQQLARGDSAGEPLEVDLEGADIPYEDIIATEI"		
BASE COUNT	380 a 522 c 554 g 236 t	
ORIGIN		
Query Match 8.9%; Score 807.2; DB 10; Length 1692;		
Best Local Similarity 99.6%; Pred. No. 2.5e-119;		
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
QY	1646	gggtccgagaaaggtgcaagtgcagtgagcccaacagccttcaagcctgtgtccacggtcagg 1705
Db	345	GGGCTCCGAGAAGGTGTCAGTGCAGGCCACACAGCCTTCAAGCCTGTGCTGCCAGGTCAGG 404
QY	1706	agccatcctgcactcctccccgagagtgcagccaccagctgcaccccgccctccaga 1765
Db	405	AGCCATCCTGCACCTCTCTCCCGGAGAGTGCCAGCCACACAGCTGCACCCGCCCTCCAGA 464
QY	1766	caagcccaagagcagagctgaagcctggcctgtgtctgtggggcgctgtcagactccgg 1825
Db	465	CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCTCAGACTCCGG 524
QY	1826	ccggaactccatgtccagctgccacacacagcagcagcagcagcagctaccagctggacc 1885
Db	525	CCGGAACCTCATGTCTCCAGCTGCCACACACAGCAGCAGCAGCTACCAAGCTGGACCC 584
QY	1886	gctgttcacacccgtgggaccacaaagccgttttgggggtctccgccacacacaccca 1945
Db	585	GCTGTGCACACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCCGCCACACATCACCCA 644
QY	1946	gggcatcgtctccaggacagcaacaatgatgagctgaagctgtgtcttctccgacgg 2005
Db	645	GGGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 704
QY	2006	aggtagcaagctggggccactgcacaaggcagacaagggccctcgtgtgtcgtccctccc 2065
Db	705	AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCTCTGTGTGTCTCCCTCCC 764
QY	2066	catctccagcagagtgcagatccagagctggaacagaagctgtgtgagaggaggagg 2125
Db	765	CATCTCCAGGACGAGTGCGAGCATCCAGGAGCTGGAGCAGAGCTGTTTGGAGAGGGAGGG 824
QY	2126	cgcctccagaagctgcagcgcagctttgagagaagagcttgctccagcctggccta 2185
Db	825	CGCCCTCCAGAAGCTGCAGCGCAGCTTTTGGAGGAAGAGGCTTGCCTTCAGGCTGGCCTA 884
QY	2186	cgaggagcggccgctgcaggaagcagctggaggcccgagagcccaagcgga 2245
Db	885	CGAGGAGCGGCCGCGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGGCCCAAGGCGCA 944
QY	2246	caagctcaagcagccttcgagaaagcagcgcgcgcagcagcagctcctgcactcaggt 2305
Db	945	CAAGCTCAAGCAGGCTCGCAGAAGACCCAGCGCGCAGCAGGCTCTGACCTGCAGT

Db	945	CAACCTCAGCAGCCCTCGCAGAAAGACCGACGCGCGCAGCAGCTCTCTGCACCTGCAAGT	1004
Qy	2306	actcagcttcacagagaaagcgcagctccgcagagagctcgcagagcctcatgaagga	2365
Db	1005	ACTCAGCTTCAGCAGCAGAGAGCGCAGCTCCGCGCAGGAGCTCGAGAGCCTCATGAAGGA	1064
Qy	2366	gcagagacctgctgagaccacgaagctcaggtctctacgagagggagagaccagctctcgccccc	2425
Db	1065	GCAGAGCTCTGCTGAGACCAAGCTCAGGTCTCTACGAGAGGGAGAGACCAAGCTTCGCGCCC	1124
Qy	2426	cgccctgagagagaccagctgagaggtgagcc	2457
Db	1125	CGCCCTGAGAGACCCAGTGGAGGTGTGCC	1156
RESULT	7		
LOCUS	AF123658	1722 bp	07-APR-1999
DEFINITION	Homo sapiens clone G3612 FEZ1 (FEZ1) mRNA, alternatively spliced, complete cds.		
ACCESSION	AF123658		
VERSION	AF123658.1	GI:4572473	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
AUTHORS	Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.		
TITLE	The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)		
MEDLINE	99199287		
REFERENCE	2 (bases 1 to 1722)		
AUTHORS	Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 2335 10th street, Philadelphia, PA 19107, USA		
FEATURES	Location/Qualifiers		
source	1..1722		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/chromosome="8"		
	/map="8p22"		
	/clone="G3612"		
	1..1722		
	/gene="FEZ1"		
	1..1722		
	/gene="FEZ1"		
	/codon_start=1		
	/product="FEZ1"		
	/protein_id="AAD23839.1"		
	/db_xref="GI:4572474"		
	/translation="MGSVSLISGHSFHSKCRASQYKLRKSHLKKLNRYSDGLLRF GFQSDGSKSSMKSGEDFFIKVQKARGSHPDYALSSDGLGQAGVDFDPST PKMLPFSNOLMGSEKGAVRPTAFKPLPSCGAILHSSPESASHQHPAPDKPKKE ELKPLCGSLSDSGNSMSLPTHSTSSYQDLPVTPVPGTSRGGSAHNTQTGIV LQDSNMSLKAISFSDGSKLGHNSKADGSCVRSPISTDECSIQEQLKLEREGA LQKLRFSFEKLSLAVPERPRRCDELEGPCKGNKLRKQASQKSRQAQVLHQA VLQQLKESRQTEYNKASEITGLKAOLKDPTRGLKLELFTQDEGALRTFKGLLELV LKAQQLKESRQTEYNKASEITGLKAOLKDPTRGLKLELFTQDEGALRTFKGLLELV ENLQPKKNAEALLRKVNLLLEQLQELRAQALARMGPPPTFPEDVPALQRELERLY WKEEKVKVYIQKQLQGSYVAMTQRNQRLEKALQQLARGDSAGEPLEVDLEGADIPYE DIATIEI"		
BASE COUNT	381 a	541 c	563 g
ORIGIN	237 t		
Query Match	8.9%	Score 807.2;	DB 10; Length 1722;
Best Local Similarity	99.6%	Pred. No. 2.5e-119;	
Matches 809; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;

```
Query Match      8.9%; Score 807.2; DB 10; Length 1722;
Best Local Similarity 99.6%; Pred. No. 2.5e-119;
Matches 809; Conservative 0; Mismatches 3; Indels 0;
```



TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
MEDLINE 99199287  
REFERENCE 2 (bases 1 to 633)  
AUTHORS Ishii,H., Baifra,K., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE Direct Submission  
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
source Location/Qualifiers  
1..633 /organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/clone="E264162"  
1..633  
/gene="FEZ1"  
1..633  
/gene="FEZ1"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="AAD33835.1"  
/db\_xref="GI:4572466"  
/translation="MGVSLSLIGSHFSKRCRASQYKLRKSSHLKLNRYSDGLLRF  
GFSODSGHKSSSKMGKSEDFYIKVSKARGSHPDYATLSSDGLGQAGVDPDPST  
PPKLPNSQNLMESEKGVNPTAFKPVLPKSGAILHSSPESASHLPAPDPKKEQ  
ELKFLGSCGALSDSGRNSMSSLPHTHSQLEPVDLEGADIPYDIATIEI"  
BASE COUNT 136 a 217 c 175 g 105 t  
ORIGIN

Query Match 3.8%; Score 346; DB 10; Length 633;  
Best Local Similarity 100.0%; Pred. No. 6.4e-46;  
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 52 atggcagcgtcagtagctctatctccggccacagcttccacagcaagcactgccggcgt 111  
Db 1 ATGGCAGCGCTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGCT 60  
QY 112 tcgcagtaacaagctgcgaagtcctctccacctcaagaagctcaaacgggtattccgacggg 171  
Db 61 TCGCAGTACAAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120  
QY 172 ctgctgaagtttggtctctccagactcgggtcagcgcaagtcagctcccaaatgggc 231  
Db 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCAAGTCCAGCTCCCAAAATGGGC 180  
QY 232 aagagcgaagacttctctacatcaaggtcagccagaagccgggctcccatcaccaca 291  
Db 181 AGACGGAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCGGGCTCCCATCACCA 240  
QY 292 gattacagcgactgccagcggggtatttagggggccagcgggtgggtgaccttgaccgc 351  
Db 241 GATTACAGGCACCTGCCAGCGGGGATTAGGGGCCAGGCTGGGTGGACTTTGACCCG 300  
QY 352 tcacacaccccccaagctcatgcccctctccaatcagctagaatgg 397  
Db 301 TCCACACCCCCCAAGCTCATGCCCTTCCCAATCATCAGTAGAATGG 346

RESULT 9  
AC025412/c AC025412 163683 bp DNA HTG 26-MAY-2000  
LOCUS Homo sapiens chromosome 1 clone RP11-30614 map 1, WORKING DRAFT  
DEFINITION SEQUENCE, 22 unordered pieces.  
ACCESSION AC025412  
VERSION AC025412.3 GI:8077107  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 163683)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 1, clone RP11-30614  
Unpublished  
2 (bases 1 to 163683)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lakocque,K., Lamazares,R., Landers,T., Lehoczkzy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 25, 2000 this sequence version replaced gi:7342167.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: I8132  
Center clone name: 306\_I\_4  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152913 bases at least Q40  
Consensus quality: 158568 bases at least Q30  
Consensus quality: 160537 bases at least Q20  
Insert size: 170000; agarose-fp  
Insert size: 161583; sum-of-contigs  
Quality coverage: 4.2 in Q20 bases; agarose-fp  
Quality coverage: 4.5 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 22 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* I 1333: contig of 1333 bp in length  
\* 1334 1433: gap of 100 bp  
\* 1434 2643: contig of 1210 bp in length  
\* 2644 2743: gap of 100 bp  
\* 2744 4634: contig of 1891 bp in length  
\* 4635 4734: gap of 100 bp  
\* 4735 7179: contig of 2445 bp in length  
\* 7180 7279: gap of 100 bp  
\* 7280 9066: contig of 1787 bp in length  
\* 9067 9166: gap of 100 bp  
\* 9167 11414: contig of 2248 bp in length



DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Melldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 5, 2000 this sequence version replaced gi:7329377.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRK

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2991

Center clone name: 8\_P\_13

----- Summary Statistics

Sequencing vector: M13; M77815; 96% of reads

Sequencing vector: Plasmid; n/a; %0.f% of reads

4.40579710144928Chemistry: dye-terminator Big Dye; 100% of

reads

Assembly program: Phrap; version 0.960731

Consensus quality: 148050 bases at least Q40

Consensus quality: 150303 bases at least Q30

Consensus quality: 151107 bases at least Q20

Insert size: 163000; agarose-fp

Insert size: 151611; sum-of-contigs

Quality coverage: 5.5 in Q20 bases; agarose-fp

Quality cov.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 10 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 8618: contig of 8618 bp in length

\* 8619 8718: gap of 100 bp

\* 8719 13037: contig of 4319 bp in length

\* 13038 13137: gap of 100 bp

\* 13138 19064: contig of 5927 bp in length

\* 19065 19164: gap of 100 bp

\* 19165 27544: contig of 8380 bp in length

\* 27545 27644: gap of 100 bp

\* 27645 36519: contig of 8875 bp in length

\* 36520 36619: gap of 100 bp

\* 36620 43866: contig of 7247 bp in length

\* 43867 43966: gap of 100 bp

\* 43967 59245: contig of 15279 bp in length

\* 59246 59345: gap of 100 bp

\* 59346 81568: contig of 22223 bp in length

\* 81569 81668: gap of 100 bp

\* 81669 110160: contig of 28492 bp in length

\* 110161 110260: gap of 100 bp

\* 110261 152511: contig of 42251 bp in length.

\* Location/Qualifiers

1..152511

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="RP11-8P13"

FEATURES

source

/clone\_lib="RPC1-11 Human Male BAC"

1..8618

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

8719..13037

/note="assembly\_fragment"

13138..19064

/note="assembly\_fragment"

19165..27544

/note="assembly\_fragment"

27645..36519

/note="assembly\_fragment"

36620..43866

/note="assembly\_fragment"

43967..59245

/note="assembly\_fragment"

59346..81568

/note="assembly\_fragment"

81669..110160

/note="assembly\_fragment"

110261..152511

/note="assembly\_fragment"

BASE COUNT 45636 a 30476 c 30973 g 44525 t 901 others

ORIGIN

Query Match 2.5%; Score 226.2; DB 60; Length 152511;

Best Local Similarity 80.7%; Pred. No. 1.8e-27;

Matches 264; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Oy 6866 ttggttcaggatattttatttatttatgacagaggtctgtctgtcccccaggtc 6925

Db 119560 TTAATTTACTCATTTCTTTTCTTTTGTGAGACAGAGTCTGCTCTGTCGCCAGGCT 119619

Oy 6926 ggagtcagtgagcatgatcatagctcactgcagctcactactcctctgggtcgaagcaatcc 6985

Db 119620 GGAGTGCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCTGTTCAAGCAATTC 119679

Oy 6986 tectctcagccctccaaactagctggagctacaggtgcgcgcacccgtgctgctgaac 7045

Db 119680 TCCTGCCCTCAGTCTCCTGAGTAGCTGGGACTACAGCGCCCACTGCTGGCTAAT 119739

Oy 7046 ttctcatTTTTTTtagggagcgggtctcgtttgttgccaaagctggtctcaacttg 7105

Db 119740 TTTTTCGTAATTTTAGTAGAGCGGGGTTTACCACATGTTAGCAGGCTGCTTTGAACCTCC 119799

Oy 7106 tggcctcaagcaatccacactgccttggcctcccaagtgctgagattgcagattgagacc 7165

Db 119800 TGACCTCAAGTGATCCACCTGCCTGGCCTCCCAAAGTCTGGGATTACAGCGGTGAGCC 119859

Oy 7166 accgtgcctggccagattttcttttta 7192

Db 119860 ACCGTGCCTGGCCATTTAATTAACCTTTTA 119886

RESULT 11

AC024740

LOCUS AC024740 179538 bp DNA HTG 07-JUL-2000

DEFINITION Homo sapiens chromosome 19 clone RP11-801D6, WORKING DRAFT

SEQUENCE, 17 unordered pieces.

AC024740

VERSION AC024740.3 GI:8570405

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 179538)

AUTHORS Waterston,R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 179538)

**AUTHORS**  
**TITLE**  
**JOURNAL**

Waterston, R.H.  
Direct Submission  
Submitted (01-MAR-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jun 17, 2000 this sequence version replaced gi:7263917.

COMMENT

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0801D06
----- Summary Statistics -----
Sequencing vector: M13, 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 97% of reads
Chemistry: Dye-terminator Big Dye; 3% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173830 bases at least Q40
Consensus quality: 175557 bases at least Q30
Consensus quality: 176446 bases at least Q20
Insert size: 242000; agarose-fp
Insert size: 177938; sum-of-contigs
Quality coverage: 5.32 in Q20 bases; agarose-fp
Quality coverage: 6.01 in Q20 bases; sum-of-contigs

```

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

*	1	1801:	contig of 1801 bp in length
*	1802	1901:	gap of unknown length
*	1902	4641:	contig of 2740 bp in length
*	4642	4741:	gap of unknown length
*	4742	4767:	contig of 4938 bp in length
*	9680	9779:	gap of unknown length
*	9780	13392:	contig of 3613 bp in length
*	13393	13492:	gap of unknown length
*	13493	17164:	contig of 3672 bp in length
*	17165	17264:	gap of unknown length
*	17265	21510:	contig of 4246 bp in length
*	21511	21610:	gap of unknown length
*	21611	27089:	contig of 5479 bp in length
*	27090	27189:	gap of unknown length
*	27190	35636:	contig of 8447 bp in length
*	35637	35736:	gap of unknown length
*	35737	45537:	contig of 9801 bp in length
*	45538	45637:	gap of unknown length
*	45638	56531:	contig of 10814 bp in length
*	56452	56551:	gap of unknown length
*	56552	66802:	contig of 10251 bp in length
*	66803	66902:	gap of unknown length
*	66903	79118:	contig of 12216 bp in length
*	79119	79218:	gap of unknown length
*	79219	94355:	contig of 15137 bp in length
*	94356	94455:	gap of unknown length
*	94456	111274:	contig of 16819 bp in length
*	111275	111374:	gap of unknown length
*	111375	130600:	contig of 19226 bp in length
*	130601	130700:	gap of unknown length
*	130701	150278:	contig of 19578 bp in length
*	150279	150378:	gap of unknown length
*	150379	179538:	contig of 29160 bp in length.

FEATURES	SOURCE
1. <b>Geographical Location:</b> The study was conducted in a rural area of a developing country, where access to healthcare services is limited.	Local Health Authorities
2. <b>Study Population:</b> The study included 150 participants, primarily women, who were selected through a random sampling method.	Community Health Workers
3. <b>Data Collection:</b> Data was collected through a series of structured interviews and focus group discussions, lasting approximately 30-45 minutes each.	Research Team
4. <b>Analysis:</b> The data was analyzed using a thematic analysis approach, identifying key themes related to the research objectives.	Statistical Software
5. <b>Ethical Approval:</b> The study was approved by the local ethics committee, ensuring that all participants provided informed consent.	Ethical Review Board
6. <b>Limitations:</b> The study has several limitations, including a small sample size and a lack of generalizability to other populations.	Research Team
7. <b>Conclusion:</b> The study concludes that there is a significant need for improved healthcare services in the study area, particularly for women.	Research Team

```
1. .179538
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="19"
```

misc_feature	1. .1801 /clone="RP11-801D6"
misc_feature	/note="assembly_name:Contig13"
misc_feature	1902. .4641 /note="assembly_name:Contig14"
misc_feature	4742. .9679 /note="assembly_name:Contig15"
misc_feature	9780. .13392 /note="assembly_name:Contig16"
misc_feature	13493. .17164 /note="assembly_name:Contig17"
misc_feature	17265. .21510 /note="assembly_name:Contig18"
misc_feature	21611. .27089 /note="assembly_name:Contig19"
misc_feature	27190. .35636 /note="assembly_name:Contig20"
misc_feature	35737. .45537 /note="assembly_name:Contig21"
misc_feature	45638. .56451 /note="assembly_name:Contig22"
misc_feature	56552. .66802 /note="assembly_name:Contig23"
misc_feature	66903. .79118 /note="assembly_name:Contig24"
misc_feature	79219. .94355 /note="assembly_name:Contig25"
misc_feature	94456. .111274 /note="assembly_name:Contig26"
misc_feature	111375. .130600 /note="assembly_name:Contig27"
misc_feature	130701. .150278 /note="assembly_name:Contig28"
misc_feature	clone_end:sp6 vector_side:left"
misc_feature	150379. .179538

BASE COUNT	42230 a	47023 c	46289 g	42392 t	1604 others
ORIGIN					

Query Match 2.5%; Score 225.6; DB 63; Length 179538;  
Best Local Similarity 67.4%; Pred. NO. 2.2e-27;  
Matches 318; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

[illegible]



```
/note="FRAM repeat: matches 1. .175 of consensus"
17527. .17670
/note="LTR16C repeat: matches 242. .387 of consensus"
complement(17559. .18000)
/note="match: GSS: Em:AQ349609"
complement(17617. .18004)
/note="match: GSS: Em:AQ350711"
18109. .18565
/note="match: STS: Em:HS812D7T"
18262. .18574
/note="AluX repeat: matches 1. .312 of consensus"
18660. .18721
/note="MIR repeat: matches 92. .147 of consensus"
19233. .19535
/note="AluSg repeat: matches 1. .303 of consensus"
19911. .20016
/note="MIR repeat: matches 47. .152 of consensus"
20023. .20252
/note="MIR repeat: matches 5. .249 of consensus"
complement(20661. .21112)
/note="match: GSS: Em:AQ019396"
complement(20698. .21115)
/note="match: GSS: Em:AQ178024"
20940. .21166
/note="MIR repeat: matches 25. .249 of consensus"
21105. .21680
/note="match: STS: Em:HS156A22T"
22012. .22288
/note="AluDb repeat: matches 13. .290 of consensus"
22356. .22515
/note="MIR repeat: matches 48. .221 of consensus"
22508. .22708
/note="L2 repeat: matches 2223. .2435 of consensus"
complement(22577. .22968)
/note="match: GSS: Em:AQ508829"
22977. .23573
/note="match: GSS: Em:AQ281488"
22998. .23126
/note="L2 repeat: matches 2157. .2280 of consensus"
23152. .23316
/note="L2 repeat: matches 2371. .2537 of consensus"
23399. .23864
/note="L2 repeat: matches 2250. .2750 of consensus"
24480. .24558
/note="MT1B repeat: matches 318. .390 of consensus"
24560. .24848
/note="AluJb repeat: matches 4. .291 of consensus"
24868. .25158
/note="AluSx repeat: matches 14. .304 of consensus"
25171. .25326
/note="FAM repeat: matches 5. .174 of consensus"
25621. .25684
/note="MIR repeat: matches 94. .157 of consensus"
26144. .26233
/note="MIR repeat: matches 50. .140 of consensus"
26337. .26421
/note="MIR repeat: matches 66. .139 of consensus"
26422. .26719
/note="AluSx repeat: matches 1. .298 of consensus"
26720. .26763
/note="MIR repeat: matches 139. .185 of consensus"

Query Match 2.5%; Score 224.6; DB 34; Length 159440;
Best Local Similarity 77.3%; Pred. No. 3.2e-27;
Matches 272; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 6878 ttatttttaatttatgacagggtctgtctgtctccccaggctgagtgcatg 6937
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127857 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
6938 catgatacatgctacgtcagcctactctctgggctcaagcaatctctgcctcagc 6997
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127797 CATAACTTGGCTTACTGCAACCTCTGCCCTCTGGGCTCAAGCAATCTCCACCTCAGC 127738
```

```
Qy 6998 attcaactagctgggactacaggtgcgcgcacgcgtgctggtacactttt 7057
||| ||| ||||| ||||| ||||| ||| ||| ||||| ||||| ||| |||||
Db 127737 CTCCCGAGTAGCTGGGACTACAGGCGCACACTACCATGCTGGCTAATGTTTGTGTTTT 127678

Qy 7058 ttgtaggacggggtctcgctttgttgccaaagctggtctcaaaacttggcctcaagca 7117
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127677 TGGTAGAGACAAGGTTTGTCTATGTTGCCAACCTGCTTGAACCTCCCTGGGCTCAAGCA 127618

Qy 7118 atccacctccctggcctcccaagtgcgagattcagatgagccaccgtgcctg 7177
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127617 GTCGTGCTGCCGCGGCTCCCAAGTGC TGGATTACAGCGGTGACCCACCATGCTGGC 127558

Qy 7178 cagattttctttattctctctctttctttctttctttctttctttctttcttt 7228
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127557 TGTAGTTATTAACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATT 127507

RESULT 13
AC016695
LOCUS
DEFINITION
AC016695 Homo sapiens chromosome 7 clone RP11-126N23, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
AC016695
VERSION
AC016695.8 GI:10645615
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 177560)
AUTHORS
Waterston,R.H.
TITLE
The sequence of Homo sapiens clone
JOURNAL
Unpublished
2 (bases 1 to 177560)
AUTHORS
Waterston,R.H.
TITLE
Direct Submission
JOURNAL
Submitted (04-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Oct 5, 2000 this sequence version replaced gi:9800603.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0126N23
----- Summary Statistics -----
Sequencing vector: M13; 72%
Chemistry: Dye-primer ET; 72% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168896 bases at least Q40
Consensus quality: 171731 bases at least Q30
Consensus quality: 173249 bases at least Q20
Insert size: 186Kb; agarose-fp
Insert size: 176260; sum-of-contigs
Quality coverage: 5.15 in Q20 bases; agarose-fp
Quality coverage: 5.49 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2804: contig of 2804 bp in length
* 2805 2904: gap of unknown length
* 2905 5617: contig of 2713 bp in length
```









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:21:50 ; Search time 4648.59 Seconds  
(without alignments)  
13639.351 Million cell updates/sec

Title: US-09-513-888-1  
Perfect score: 9048  
Sequence: 1 gccttccaagacctgcc.....tgccattctcagccctct 9048

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues  
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: em\_estba:\*  
42: em\_estfun:\*  
43: em\_esthum1:\*  
44: em\_esthum2:\*  
45: em\_esthum3:\*  
46: em\_esthum4:\*  
47: em\_esthum5:\*  
48: em\_esthum6:\*  
49: em\_esthum7:\*  
50: em\_esthum8:\*  
51: em\_esthum9:\*  
52: em\_esthum10:\*  
53: em\_esthum11:\*  
54: em\_esthum12:\*  
55: em\_esthum13:\*  
56: em\_esthum14:\*  
57: em\_esthum15:\*  
58: em\_esthum16:\*  
59: em\_esthum17:\*  
60: em\_esthum18:\*  
61: em\_esthum19:\*  
62: em\_esthum20:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estov1:\*  
68: em\_estov2:\*  
69: em\_estpl1:\*  
70: em\_estpl2:\*  
71: em\_estpl3:\*  
72: em\_estpl4:\*  
73: em\_estpl5:\*  
74: em\_estro1:\*  
75: em\_estro2:\*  
76: em\_estro3:\*  
77: em\_estro4:\*  
78: em\_estro5:\*  
79: em\_estro6:\*  
80: em\_estro7:\*  
81: em\_estro8:\*  
82: em\_estro9:\*  
83: em\_estro10:\*  
84: em\_estro11:\*  
85: em\_estro12:\*  
86: em\_estro13:\*  
87: gb\_est41:\*  
88: gb\_est42:\*  
89: gb\_est43:\*  
90: gb\_est44:\*  
91: gb\_est45:\*  
92: gb\_est46:\*  
93: gb\_est47:\*  
94: gb\_est48:\*  
95: gb\_est49:\*  
96: gb\_est50:\*  
97: gb\_est51:\*  
98: gb\_est52:\*  
99: gb\_est53:\*  
100: gb\_est54:\*  
101: gb\_est55:\*  
102: gb\_est56:\*  
103: gb\_est57:\*  
104: gb\_est67:\*  
105: gb\_est68:\*  
106: gb\_est69:\*  
107: gb\_est70:\*  
108: gb\_est71:\*  
109: gb\_est72:\*  
110: gb\_est73:\*  
111: gb\_est74:\*  
112: em\_esthum21:\*  
113: em\_esthum22:\*  
114: em\_esthum23:\*  
115: em\_estom1:\*  
116: em\_estom2:\*

117: em\_estpl6:\*  
118: em\_estpl7:\*  
119: em\_estpl8:\*  
120: em\_estpl9:\*  
121: em\_estpl10:\*  
122: em\_estpl11:\*  
123: em\_estpl12:\*  
124: em\_estpl13:\*  
125: em\_estpl14:\*  
126: em\_estpl15:\*  
127: em\_estpl16:\*  
128: em\_estpl17:\*  
129: em\_estpl18:\*  
130: em\_estpl19:\*  
131: em\_estpl20:\*  
132: em\_estpl21:\*  
133: em\_estpl22:\*  
134: em\_estpl23:\*  
135: em\_estpl24:\*  
136: em\_estpl25:\*  
137: em\_estpl26:\*  
138: em\_estpl27:\*  
139: em\_estpl28:\*  
140: em\_estpl29:\*  
141: em\_estpl30:\*  
142: em\_estpl31:\*  
143: em\_estpl32:\*  
144: em\_estpl33:\*  
145: em\_estpl34:\*  
146: em\_estpl35:\*  
147: em\_estpl36:\*  
148: em\_estpl37:\*  
149: em\_estpl38:\*  
150: em\_estpl39:\*  
151: em\_estpl40:\*  
152: em\_estpl41:\*  
153: em\_estpl42:\*  
154: em\_estpl43:\*  
155: em\_estpl44:\*  
156: em\_estpl45:\*  
157: em\_estpl46:\*  
158: em\_estpl47:\*  
159: em\_estpl48:\*  
160: em\_estpl49:\*  
161: em\_estpl50:\*  
162: em\_estpl51:\*  
163: em\_estpl52:\*  
164: em\_estpl53:\*  
165: em\_estpl54:\*  
166: em\_estpl55:\*  
167: em\_estpl56:\*  
168: em\_estpl57:\*  
169: em\_estpl58:\*  
170: em\_estpl59:\*  
171: em\_estpl60:\*  
172: em\_estpl61:\*  
173: em\_estpl62:\*  
174: em\_estpl63:\*  
175: em\_estpl64:\*  
176: em\_estpl65:\*  
177: em\_estpl66:\*  
178: em\_estpl67:\*  
179: em\_estpl68:\*  
180: em\_estpl69:\*  
181: em\_estpl70:\*  
182: em\_estpl71:\*  
183: em\_estpl72:\*  
184: em\_estpl73:\*  
185: em\_estpl74:\*  
186: em\_estpl75:\*  
187: em\_estpl76:\*  
188: em\_estpl77:\*  
189: em\_estpl78:\*

190: gb\_gss25:\*  
191: gb\_gss26:\*  
192: gb\_gss27:\*  
193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	704.2	7.8	705	28	AL118597	AL118597 DKFZp761D
2	549.8	6.1	599	142	N31948	N31948 YY22g12.s1
3	525.2	5.8	547	40	AW137540	AW137540 UI-H-B11-
4	524.4	5.8	545	38	AW007737	AW007737 wt68e06.x
5	521.2	5.8	529	92	AW590735	AW590735 hg48h03.x
6	502.4	5.6	619	142	N21184	N21184 yx41a10.s1
7	484.4	5.4	817	27	AI984777	AI984777 wr85b07.x
8	480.8	5.3	743	107	BE410921	BE410921 601303579
9	468.8	5.2	472	28	AL135575	AL135575 DKFZp762B
10	468.4	5.2	487	19	AI357233	AI357233 qx63a04.x
11	468	5.2	673	107	BE384131	BE384131 601272956
12	447.2	4.9	452	39	AW051759	AW051759 wx28h07.x
13	445.4	4.9	447	15	AI042490	AI042490 ox62e04.x
14	440.4	4.9	442	19	AI360882	AI360882 qy01c12.x
15	439.8	4.9	453	23	AI633596	AI633596 th70a03.x
16	439.4	4.9	451	23	AI636674	AI636674 ts92e04.x
17	433.6	4.8	643	142	N35845	N35845 yx89f10.r1
18	427.8	4.7	432	38	AW002410	AW002410 wu61f05.x
19	423.4	4.7	425	134	BE049448	BE049448 wx86f10.x
20	421.4	4.7	434	22	AI623326	AI623326 ts80d08.x
21	416.4	4.6	418	15	AI078630	AI078630 ox51b09.x
22	403.6	4.5	495	142	N42784	N42784 yy22g12.r1
23	401.4	4.4	796	106	BE279166	BE279166 601156630
24	392.4	4.3	450	23	AI652496	AI652496 wb29h02.x
25	391.8	4.3	401	19	AI362152	AI362152 qy44h07.x
26	387	4.3	558	107	BE384676	BE384676 601276954
27	385	4.3	461	142	N24642	N24642 yx89f10.s1
28	368.8	4.1	531	106	BE276168	BE276168 601144424
29	365	4.0	397	39	AW028197	AW028197 wv84a01.x
30	362.4	4.0	797	109	BE540725	BE540725 601085002
31	351.4	3.9	463	1	AA020852	AA020852 ze64b08.r
32	342.6	3.8	471	141	H09757	H09757 ym01d02.s1
33	342.6	3.8	751	106	BE312985	BE312985 601150246
34	336.6	3.7	344	38	AW016544	AW016544 UI-H-B10p
35	335	3.7	357	15	AI080440	AI080440 ox82c12.s
36	303	3.3	328	141	H09841	H09841 ym01d02.r1
37	300	3.3	323	23	AI636335	AI636335 tz78a05.x
38	299.4	3.3	302	140	F11822	F11822 HSC32A091.n
39	299.2	3.3	349	145	T65388	T65388 yc73b09.s1
40	299.2	3.3	448	144	R88008	R88008 ym87b02.r1
41	296.2	3.3	326	28	AL047147	AL047147 DKFZp586E
42	296	3.3	296	28	AL134288	AL134288 DKFZp347J
43	279	3.1	347	144	R87966	R87966 ym87b02.s1
44	257.4	2.8	274	140	F09471	F09471 HSC32A092.n
45	254.8	2.8	407	38	AV663617	AV663617 AV663617

## ALIGNMENTS

RESULT 1  
AL118597  
LOCUS AL118597 705 bp mRNA 29-FEB-2000  
DEFINITION DKFZp761D0110.r1 761 (synonym: hamy2) Homo sapiens cDNA clone  
ACCESSION AL118597  
VERSION AL118597.1 GI:5924496  
KEYWORDS EST.  
SOURCE human.

ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 (bases 1 to 705)									
TITLE	Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.									
COMMENT	EST (Blum, et al.) Unpublished (1999) Contact: Blum H MIPS									
FEATURES	An Klopferspitz 18a D-82152 Martinsried, Germany This is the 5' sequence of the clone insert Research Center (DKFZ); Email s.wiemann@kfz- heidelberg.de; sequenced by LMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available. This clone (DKFzp761D0110) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. Location/Qualifiers 1..705 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="DKFzp761D0110" /clone_lib="761 (synonym: hamy2)" /tissue_type="amygdala" /dev_stage="adult" /lab_host="DH108" /note="vector: pSport1; Site_1: NotI; Site_2: SalI"									
BASE COUNT	148 a	205 c	203 g	149 t						
ORIGIN										
Query Match	7.88; Score 704.2; DB 28; Length 705;									
Best Local Similarity	99.7%; Pred. No. 2.1e-61;									
Matches	703; Conservative	2; Mismatches	0; Indels	0; Gaps	0;					
Qy	6144	ccagtgccgggtgatgccagccctgggcccagccctgttactgggtcttgcgaat	6203							
Db	1	CCATGGCGGGTGATGCCAGCCCTGGGGCCAGCCCTGTTACTGGTCTTGCAGAAAT	60							
Qy	6204	gggagctgagcagcctctgacagccagtcagctttgacctcggtgaccactctcttta	6263							
Db	61	GGGAGCTGAGCAGCCTCTGGACAGCCAGTACCTTTGACCTCGCTGACCACCTCTCTTTA	120							
Qy	6264	agccatagaccctgagccctgggctgggtgctgggaaggagggttgaaaccaccgtga	6323							
Db	121	AGCCATAGACCTCAGAGCCCTGGGCTGGGTGCTGGGAAGGGAGGGTTGAACACCACTGA	180							
Qy	6324	accagagggtggtgctttccagkcaacctcagggagcctccccatctgtccagctggggc	6383							
Db	181	ACCAGAGGCTGGCTTTCCAGTCAACCTCAGGGAGGCTCCCCATCTGTCAGACTGGGGC	240							
Qy	6384	cagaggctgggagtcctacctgcttcacgttgggccggcgactactcggaatgttttc	6443							
Db	241	CAGAGGCTGGGAGTCCCTACCTGCTCAGGTGGCCGGCGGCTACTCTGGAATGTTTTTC	300							
Qy	6444	cctcccagaatacaagcttttgctgatccagaagagcccatatcaactaagatggccatat	6503							
Db	301	CCTCCCAGAAATCAAGCTTTTGTTGATCCAGAGAGGCCATATCAGTAAGATGGCATAT	360							
Qy	6504	atgtgactgggcatcttccctcctctcctacagccaggttttagcgcaaacctttccc	6563							
Db	361	ATGTGATCTGGGCATTTTCTCTCTGCTACAGCCAGGTTTACGGCGAACCCTTTCCCC	420							
Qy	6564	cttagcaccttcagggtcagttctcctctctctctctctctctctctctctctctctctc	6623							
Db	421	CTTAGCACCTTCAGGGCTGAGTTCTGGGTTTCTAGAGGTCAGGAGCGCTCCTCTCAGAGCGC	480							
Qy	6624	cagaaagccagagccccaaagcaggaacaaaaaggagcatacacacagcagtgatgaatgc	6683							
Db	481	CAGGAAGCCAGAGCCCAAGCAGGACCAAAAGAGGATACACACAGCAGTGTCAATAGC	540							

```
QY 8446 tcaccaagatcgagctc-tgcccctggacctgggaagactggactgtcaagggttcc 8504
Db 599 TCACCAAGATCGGACTNTTGGCCCTTGGACCTGGGAACGACTGGACTGFCAC -GGGTTCC 541
QY 8505 ctccctagctctccagtgaaactccttcgcaaggcacacacagccctatagcaactgagctca 8564
Db 540 NTCCCTAGCTCTCCCAAGTGAACCTCTGCCAGGACACACAGCCCTTATAGCACTGAGCTCA 481
QY 8565 catggagctgggatagggggcatctctcccccagagagcactcagtgagcctcctgtg 8624
Db 480 CATGGGACTGGGATATGGGGGATCTCTTCCCAGAGAGGCACTCAGTGAGGCTCTCTGTG 421
QY 8625 cctggcccagctcgtggccatctcttagtgagacagtggtgcccgaact-aagccaggcc 8683
Db 420 CCTGGCCCCAGCCTGGGCCATCTCTTAGGTGAGACAGATTGCCCCGAAACTAAAGCCAGGCC 361
QY 8684 tggctggagagcagcagcttgggagagggatttcctcgcagacctcaagccatcgtc 8743
Db 360 TGGCTGGAGGAGCAGCAGCTTGGGAGAGGATTTCCCTGCGAGAGCTCAAGCCATCATGC 301
QY 8744 ggtgggtgctgccatgacagagctgcacccctggccagcgggctgctcaccacctc 8803
Db 300 GTGGGTGCTGCCATGACAGAGCTGCACCCCTGGCCACAGCGGCTGCTCACCCACCTC 241
QY 8804 ttgtgcaagtggtcccttgtgctgcgctgcaggcagagctggagcccccagcagagga 8863
Db 240 TTGTGCAAGGTGGCCTTTTGTGTTGCGCCTGCGAGGAGCTGGAGCCCCCAGCAGAGGCA 181
QY 8864 ggtcgggagcagcagcagctcctggaagatgacatgattttcttctgtgtgttctt 8923
Db 180 GGTGGGAGCGGACGACGACTCTGGAAGATGTACATGTTATTTTCTTCTTGTGGTTCTT 121
QY 8924 gtttggttggttcttcttgacagctcattttatttttgcagtcacttttggccatg 8983
Db 120 GTTGTGTTGGTTGCTTTTGACAGCTCATTTTATTTTGGCGTCACTTTTGGCCATG 61
QY 8984 taaactatttggcaattttatgtttttattttatgaataaagaatgccatttctcaagc 9043
Db 60 TAAACTATTGTGGCAATTTTATGTTTTTATTATGATTAAGAATGCCATTTTCTCACGC 1

RESULT 3
LOCUS AW137540/c
DEFINITION UI-H-B11-acv-h-12-0-UI-s1 NCI_CGAP_Sub3 Homo sapiens CDNA clone
IMAGE:2715887 3', mRNA sequence.
ACCESSION AW137540
VERSION AW137540.1 GI:6141858
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: M3 Forward
POLYA=yes.
Location/Qualifiers
1. .547
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:2715887"
/lab.lib="NCI_CGAP_Sub3"
/lab.host="DH10B (Life Technologies)"
/note="vector: pYT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub3 library is a subtracted library derived from
the NCI_CGAP Sub1 library, which is a subtracted library
derived from BI. BI constitutes a mixture of 21
normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775
, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342
, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831,
1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991,
1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,
1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids
985608-986759, 1101192-1101959, 1217928-1220615);
NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
Clonoids 1057416-1061255, 1144584-1145351). Subtraction
was performed as previously described [Bonaldo, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG_LIB=NCI_CGAP_Brn23
TAG_TISSUE=brain
TAG_SEQ=ATATC".

BASE COUNT 143 a 147 c 138 g 119 t
ORIGIN

Query Match Similarity 5.8%; Score 525.2; DB 40; Length 547;
Best Local Similarity 99.4%; Pred. No. 1.le-43;
Matches 527; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8513 tctccagtgaaactcctccagagagagcactcagtcagtcagtcacatggagc 8572
Db 547 TCTCCCCGTGAACCTCGTCCAGGCACACAGCCCTATAGCACTGAGCTCATGTGGGAC 488
QY 8573 tgggtatgggggcatctcttcccagagagcactcagtcagtcagtcctgtgctggccc 8632
Db 487 TGGGATATGGGGGCACTCTTCCCCAGAGAGGCACTCAGTGAGCCCTCTGTGCTGGCCC 428
QY 8633 cagtcctggccatctcttagtgagacagtggtgcccgaactaagccaggtcgtgtgag 8692
Db 427 CAGCCTGGGCCATCTCTTAGGTGAGACAGTGTGCCGAAACTAAGCCAGGCTGTGGTGAG 368
QY 8693 gagcagcagctgggagagagatttccctcagacctcaagccatcagtggtgtgc 8752
Db 367 GAGCAGAGCTTGGGAGAGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGC 308
QY 8753 tgcctgacagaggtgcacccctggccagcggtgtgctcacccacctcttctgtgcaag 8812
Db 307 TGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGTCAG 248
QY 8813 gtggccttctgtcgcctgcaggcagagctggagccccccagcagagcaggtcggagc 8872
Db 247 GTGGCTTTGTGCTGCGCTGCGAGCAGAGCTGGAGCGCCCGCCAGCAGAGCGAGCTGGGAC 188
QY 8873 ggaccagcatctggaagatgtacatagttattttctcttctgtgttcttctgttggttt 8932
```

Db	187	GCACGAGCATCTGGAAGATGACATAGTATTTTCCTTTGTGGTTCCTTTGTGGTTT	128
Qy	8933	ggttgctctttgcagagcttcattttatttttgacgtcaccttttggccatgtaaacatt	8992
Db	127	GGTTGTCTTTTGACAGCTTCATTTTATTTTTCAGTCACCTTTTGGCCATGTAAACTATT	68
Qy	8993	tgtggcaatttatgttttttttttttatgaataaagaatgccatttctcagc	9042
Db	67	TGTGCAATTTATGTCTTTTATGAATAAAGAATGCCATTTCTCAGC	18
RESULT 4			
AW007737/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
* Query Match			
Best Local Similarity 99.8%; Score 524.4; DB 38; Length 545;			
Matches 525; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1646	gggtccgagaggtgcagtgagggccacagcttcaagcctgtgtgccacggtcagg	1705
Db	526	GGGTCCGAGAAAGGTCAGTGAAGCCACAGCTTCAAGCTGTGTGCCACGGTCAGG	467
Qy	1706	agccatctgcactctctcccgagagtgccagccacacagctgcaccccgccctccaga	1765
Db	466	AGCATCTCTGCATCTCTCCCGGAGAGTGCAGCACACAGCTGCACCCGCCCTCCAGA	407
Qy	1766	caagcccaagagcagagctgaagcctggcctgtgtctggggcgctgtcagactcgg	1825





Db 76 GTCACTTTTGGCCATGTAACTATTGTGGCAATTTTATGTATTTTATTTATGAATAAAG 17

QY 9027 aatgccatttccag 9042  
|||||

Db 16 AATGCCATTTCACG 1

RESULT 7  
AI984777/c mRNA EST 08-MAR-2000  
LOCUS AI984777  
DEFINITION wr85b07.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2494453 3' similar to TR:060299 O60299 KIAA0552 PROTEIN. ;, mRNA sequence.  
ACCESSION AI984777  
VERSION AI984777.1 GI:5812054  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 817)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbrrp/image/image.html  
Insert length: 1307 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 428.

FEATURES  
Source  
1..817  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2494453"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI\_CGAP\_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 133 a 235 c 266 g 180 t 3 others  
ORIGIN

Query Match 5.4%; Score 484.4; DB 27; Length 817;  
Best Local Similarity 97.2%; Pred. No. 1e-39;  
Matches 514; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1646 gggctccgagaggggtcagtgaggccccacagccttaagcctgtgtctgccaggtccag 1705  
|||||

Db 529 GGGCTCGTAGAGGGTGCAGTAAGCCACACAGCTTCAAGCTTGCTGCCACGGTCAGG 470  
|||||

QY 1706 agccatctgcactctccccggagagtgccagccacagctgcacccgccccctccaga 1765  
|||||

Db 469 AGCCATCCCTGCATCTCCCGGAGAGTGCCAGCCACACAGCTGCACCCGCCNTCCAGA 410  
|||||

QY 1766 caagcccaaggagcagagctgaagcctgtgtctgtggcgctgtcagactcccg 1825  
|||||

Db 409 CAAAGCCAAAGGAGCAGAAAGCTGAAGCCTGCCTGTGCTCTGTGGGGCGCTGTACAGACTCCGG 350

QY 1826 ccggaactcatgtccagcctccacacacag-caccacagcagcagctaccagct--gga 1882  
|||||

Db 349 CCGGAATCTCATGTCCAGCCTGCCACACACACGCCCCACAGAAAGTTACCAAGTTGGGA 290  
|||||

QY 1883 ccgcgtggtcacacaccgctgggaccccaagcgttttggggctccgcccccaacaatcac 1942  
|||||

Db 289 CCGCCTGGTCCACACCCGCTGGGACCCACAAAGCCGTTTTGGGGCTCCGCCACAAATCAC 230

QY 1943 ccaggcatcgtctctccaggacagcaaatgatgagcctgaagcctctgtctcttccga 2002  
|||||

Db 229 CCAGGCAATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGA 170  
|||||

QY 2003 cggaggtagcaagctggccactcgaacaggcagacaaaggccctcgtgtctccgctc 2062  
|||||

Db 169 CGGAGTAGCAAGCTGGGCCACTCGAACAAAGGCGACAAAGGGCCCTCGTGTCTCGCTC 110  
|||||

QY 2063 ccccatctccaggcagtgagtcagcatccaggagctggaaacagaagctgttgagagggga 2122  
|||||

Db 109 CCCATCTCCACGGAGTAGTCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGA 50  
|||||

QY 2123 gggcgccctccagaaactcagcagcagcgtttgaggagaaggagccttgc 2171  
|||||

Db 49 GGGCGCCCTCCAGAACTGCAGCGCAGCTTTGAGAGAAAGGAGCTTGCC 1

RESULT 8  
BE410921/c mRNA EST 21-JUL-2000  
LOCUS 601303579F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3638203 5', mRNA sequence.  
ACCESSION BE410921  
VERSION BE410921.1 GI:9347371  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 743)  
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM338 row: m column: 20  
High quality sequence start: 49  
High quality sequence stop: 688.

FEATURES  
Location/Qualifiers  
1..743  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3638203"  
/clone\_lib="NIH\_MGC\_21"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 119 a 229 c 226 g 168 t 1 others  
ORIGIN

```
Query Match          5.3%; Score 480.8; DB 107; Length 743;
Best Local Similarity 81.1%; Pred. No. 2.4e-39;
Matches 658; Conservative 0; Mismatches 13; Indels 140; Gaps 3;

QY 1923 ggctccgcccacacatcacccaggggcctcctccaggacagcaacatgatgagcctg 1982
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 711 GGGCGCGCCCAACACACAGG-CATCGTCTCCAGGACGACGCAACATGATGAGCCTG 653

QY 1983 aaggctctctctctccagcagaggttagcaagctgggcccactcgaacaagcagacaag 2042
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 652 AA-GCTCTCTCTCTCCAGCAGGAGGTAGCAAGCTGGGCCACTGGAACAAGGACAGCAAG 594

QY 2043 ggccctctgtctcgtcccccctcctccagcgagcagctcagcagcagcagcagcagc 2102
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 593 GGCCCTCTGTGTCTGCTGCCCTCCATCGCACGGAGGTGAGCATCCAGGAGCTGGAG 534

QY 2103 cagaagctgttgagaggaggcgcctccagaagctgcagcgcagcgtttgaggagaag 2162
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 CAGAAGCTGTTGAGAGGGAGGGCGCCCTCCAGAAAGCTGCAGCGAGCTTTGAGGAGAG 474

QY 2163 gagctgtccagcctgcctctacagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2222
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 GAGCTTGCTCCAGCTGCGCTTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAG 414

QY 2223 ggcccgagcccaaggcgcgaacaagctcaagcagcagcgcgcgcgcgcgcgcgcgcgc 2282
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 GGCCCGAGCCCAAGGCGGCAACAGCTCAA----- 380

QY 2283 cagcaggtcctgcacctgcagcttgcagcttcagcagagagaagcgcgcgcgcgcgcag 2342
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 ----- 380

QY 2343 gagctcgagcctcatgaagagcagcagcctgctggagaccaagctcaggtccctacgag 2402
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 -----GTCTTACGAG 372

QY 2403 agggagaagaccagcttcgccccgcctggagagaccagtcaggtgagggcacaac 2462
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 AGGGAGAAGACCAGCTTGCGCCCCCGCTGGAGGAGACCCAGTGGAGGTGAGGCCACAC 312

QY 2463 agggctcatgggttggttggtcagcgggttggtgcgcgcgcgcgcgcgcgcgcgcgcgc 2522
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 AGGGCTCATGGGTTTGGGTGGTGCAGCGGTTTGGCGCCAGTACCCCTCTCTCTGTGTG 252

QY 2523 ctggccaatagcgtgcaaacacagaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2582
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 CCGGCCAATAGCGTGCAACACAGACACCGCGCAGGCAAGCGGGGCTTAATGTGTGCTTTA 192

QY 2583 tcaccacaagaggcctccctgcacaaccatgttggggatcgactcatctcagcttc 2642
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 TCACCCAAAGAGGGGCTCCCTGCAAAACCATGTTGGGGATCGACTTACATCTGAGCTTC 132

QY 2643 ctctgtcccccaccatcacctctcagcctcagcttgcagtttcccaagtgcacatta 2702
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 CTCCTGTCCCCACCATCACCTTCCCTGCTTCTAGATTTCAGTTTCCCAAGTGGCCATT 72

QY 2703 aatcatgaagccggaagccagatgaccaag 2733
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 AATCATGAAGCCGGAAGCCAGATGTGCCAAG 41
```

```
RESULT 9
AL135575/c  AL135575 472 bp mRNA EST 25-FEB-2000
LOCUS      DKEZp762B204_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
DEFINITION DKEZp762B204 5', mRNA sequence.
ACCESSION  AL135575
VERSION    AL135575.1 GI:6603762
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 472)  
Koehler,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.  
EST (Koehler,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)  
Unpublished (1999)  
Contact: Koehler K  
MIPS  
Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the Charite,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
No sl sequence available.  
This clone (DKFzp762B204) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source  
1..472  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFzp762B204"  
/clone\_lib="762 (synonym: hmel2)"  
/tissue\_type="melanoma (MeWo cell line)"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="vector: pSport1; Site\_1: NotI; Site\_2: SalI"  
BASE COUNT 133 a 134 c 117 g 88 t  
ORIGIN

Query Match 5.2%; Score 468.8; DB 28; Length 472;  
Best Local Similarity 99.6%; Pred. NO. 4.6e-38;  
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 8574 gggataggggcacatcttccccagagagcactcagtcagcctcctcctgctgctgcccc 8633
Db 472 GGGATATGGGGCATCTCTTCCCCAGAGAGGCACCTAGTGAGCCTCTGTGCTGGCCCC 413

QY 8634 agctgggccatctcttaggtgagacagttgcccaaaacagcagcgcctggctggag 8693
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 AGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCTGGCTGGAGG 353

QY 8694 agcagcagctgggagaggggatttccctgcagacccaagcacaatcgcggtgct 8753
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 AGCAGCAGCTTGGGAGAGGGATTTCCTCTGCAGAGCTCAAGCCATCATCGGTGGTGCT 293

QY 8754 gccatgacagagctgcacccctggccagcgggctcgcacccacaccttctgtcaagg 8813
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 GCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTACCCACCTCTTGTGCAAGG 233

QY 8814 tggccttctgctgcgcctgcaggcagagctgagccccccagcagcagcagcagcagc 8873
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 TGGCCTTTGTGCTGCGCCTGCAGGCGAGAGCTGGAGCCCCCAGCAGCAGCGCTGGAGG 173

QY 8874 gaccagcactggaagatgtacatagttatttctcttctgtgttcttctgtgtgtgtg 8933
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 CACCAGCATCTGGAAGATGTACATAGTTATTTTCTCTTGTGTGTCTTGTGTGTG 113

QY 8934 gtttgcctttgacagcttcatttattttgacgcacacttttggccatgtaaacattt 8993
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 GTTTCCTTTGACAGTTTCATTTTATTTTGACGTCACCTTTTGGCCATGTAACATATT 53

QY 8994 gtggcaatttatgtttttatttatgaataaagaatgccatttctcagccc 9045
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 52 GTGGCAATTTATGTTTTTATTTATGAATAAAGATGCCATTTCTCACGCC 1
```

RESULT 10  
AL1357233/c AL1357233 487 bp mRNA EST 15-FEB-1999  
LOCUS

DEFINITION	qx63a04.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:2005998 3', mRNA sequence.
ACCESSION	AI357233
VERSION	AI357233.1 GI:4108854
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 487)
TITLE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://www.bio.llnl.gov/bbrp/image/image.html">www.bio.llnl.gov/bbrp/image/image.html</a> Insert length: 743 Std Error: 0.00 Seq primer: -400P from Gibco High quality sequence stop: 444. Location/Qualifiers 1..487 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2005998" /clone_lib="NCI_CGAP_GC4" /tissue_type="pooled germ cell tumors" /lab_host="DH10B" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
FEATURES	
SOURCE	133 a 134 c 116 g 104 t
BASE COUNT	
ORIGIN	
Query Match	5.2%; Score 468.4; DB 19; Length 487;
Best Local Similarity	99.8%; Pred. No. 4.9e-38;
Matches	469; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	8573 tgggatgggggcatcttccccagagaggaactcagtgagcctctgtgctgtgccc 8632
Db	
Qy	8633 cagtctgggccatctcttagtgagacagttgcccgaactaaagccagggcctgctggag 8692
Db	
Qy	8693 gagcagcagcttggggagagggattccctcgagacctcaagccatcatcggtgggtgc 8752
Db	
Qy	8753 tgccatgacagaggtgcaccccttggccagcggggctgtcacccacctttgtgaag 8812
Db	
Qy	8813 gtggcccttctgtctgccttcgagcagactgagccccagcagagcaggtggagac 8872
Db	
Qy	8872 gtggcccttctgtctgccttcgagcagactgagccccagcagagcaggtggagac 188
Db	
Qy	8873 ggaccagcatctggaagatgtacatagttattttctcttctgtggtttctgttggttt 8932
Db	

Db	187	GGACAGCATCTGGAAGATGTACATAGTATATTTCTCTTTGTGGTTCTCTGTGTTGGTTT	128
Qy	8933	gnttgctttgacagcttcattttatttttgacgtcactcttttttgccatgtaaaactatt	8992
Db	127	GGTTGCTTTTGACAGCTTCATTTATTTTTCAGCTCAGTCACTTTTGGCCATGTAACATATT	68
Qy	8993	tytgcaattttatgtttttttttttatgaataaagaatgccattttctcacg	9042
Db	67	TGTGGCAATTTTATCTTTTATTTATGAATAAAGATGCCATTTCTCACG	18
RESULT	11		
BE384131/c			
LOCUS			
DEFINITION	BE384131	673 bp mRNA	EST 21-JUL-2000
	60127956F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614076		5'
ACCESSION	BE384131	mRNA sequence.	
VERSION	BE384131.1	GI:9329496	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 673)		
JOURNAL	NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC). Unpublished (1999). Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCM275 row: p column: 13 High quality sequence stop: 670. Location/Qualifiers 1..673 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3614076" /clone_lib="NIH_MGC_20" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into ECORI/XhoI sites using the following 5', adaptor: GCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)		
BASE COUNT	110 a	211 c	208 g
ORIGIN		144 t	
Query Match	5.2%;	Score 468;	DB 107; Length 673;
Best Local Similarity	81.9%;	Pred. No. 4.6e-38;	
Matches 663;	Conservative	0; Mismatches 5;	Indels 142; Gaps
Qy	1943	ccaggagcatgcctccaggacagacaacatgatgcctgaaggctctctctctccga	2002
Db	669	CCAGGCAATCGTCCYCCA-GACAGCAACATGATGAGCCTGAAGGCTCTGTCTCTCCGA	611
Qy	2003	cggaggtagcaagctggggccactcgaaacaggcagacaagg-cacctggtgtgcgcgt	2061
Db	610	CGAGGTAGCAAGCTGGGGCCATCTCGAACAGGCAGACAGGCCCCCTCTGTGTCTCGCT	551
Qy	2062	cccccatctccacgacagctgagtcagcatccagagctggacagagaagctgttggagagg	2121



Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1. (bases 1 to 447)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LLNL ; contact the  
Immunology Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.

Insert Length: 656 Std Error: 0.00

Seq primer: -40m13 fwd. RT from Amersham

High quality sequence stop: 435.

Location/Qualifiers

1. 447

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1660926"  
/clone\_lib="Soares.NHMPu\_S1"  
/tissue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;  
Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NBSM, pregnant uterus  
NHMPu, and fetal heart NBH19W) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

62 a 133 c 156 g 96 t

BASE COUNT  
ORIGIN

Query Match	4.9%	Score 445.4	DB 15	Length 447
Best Local Similarity	99.8%	Pred. No. 9.6e-36		
Matches 446	Conservative 0	Mismatches 1	Indels 0	Gaps 1
QY	1725	ccgagagatccagccaccagctcacccgcgccccctccagacaagcccaagagcagagag	1784	
Db	447	CCGAGAGTGCACGCCACCAAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAGGAG	388	
QY	1785	ctgaagcctggcctgtgctctggggcgctgtcagactccggccggaaactccatgtccagc	1844	
Db	387	CTGAAGCCTGGCCTGTGCTCTGGGGCGCGCTGTCAACATCCGGCCGGAACCTCATGTCCAGC	328	
QY	1845	ctgcccacacagcacagcagcagcagctacagctgagccgctggtcacaccccgctggga	1904	
Db	327	CTGCCACACACAGCAGCAGCAGCTACCAAGTGGACCCGCTGGTTCACACCCCGTGGGA	268	
QY	1905	cccacaagccgctttgggggctccgcccacaacatcaccacagggcatactgtctccaggaac	1964	
Db	267	CCCAACAAGCCGTTTTGGGGGCTCCGCCACAACAATACCCAGGGCATCTCTCCAGGAC	208	
QY	1965	agcaaatgatgacgcttgaaggtctgtctctctccagcagaggtagcaagctggggcac	2024	
Db	207	AGCAACATGATGAGCCCTGAAGGCTCTGTCTTCTCCACGGAGGTAGCAAGCTGGGGCAC	148	
QY	2025	tcgaacaaggcagacaagggccctctgtgtccgctcccccatctccacggagcagtgcc	2084	
Db	147	TCGAACAAGGCNACAGGGGCCCTCTGTGTCTCCGCTCCCCCATCTCCACGGACGAGTGC	88	
QY	2085	agcatccagagagctggacaagaagctgtttgagagggagggcgccctccagaagctcgag	2144	
Db	87	AGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAG	28	
QY	2145	cgcagctttgaggagaaggagcttgc	2171	
Db	27	CGCAGCTTTGAGGAGAAGGAGCTTGCC	1	

RESULT 14	
AI360882/c	
LOCUS	442 bp mRNA EST 15-FEB-1999
DEFINITION	gY01c12.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2010742 3',
ACCESSION	AI360882
VERSION	AI360882.1 GI:4112503
KEYWORDS	mRNA sequence.
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1 (bases 1 to 442)
COMMENT	NCI/NINDS-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index Unpublished (1998) Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://www-bio.llnl.gov/bbrp/image/image.html">www-bio.llnl.gov/bbrp/image/image.html</a> Insert Length: 707 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 415. Location/Qualifiers 1. .442 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2010742" /tissue_type="glioblastoma (pooled)" /lab_host="DH10B" /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTCAAGTGGGCGGCATATCTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	126 a 124 c 109 g 83 t
ORIGIN	

[illegible]







**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:52:25 ; Search time 318.41 Seconds  
(without alignments)  
10674.903 Million cell updates/sec

Title: US-09-513-888-1

Perfect score: 9048

Sequence: 1 gccttccaagaccctgccc.....tgccatttctcagccctct 9048

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_36:\*

1: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
2: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
3: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
4: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
5: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
6: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
7: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
8: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
9: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
10: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
11: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
12: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
13: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
14: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
15: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
16: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
17: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
18: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
19: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
20: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
21: /cgn2\_2/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	365.8	4.0	393	16 T23583	Human gene signatu
C 2	212.2	2.3	700	20 X30339	DNA encoding a hum
C 3	211.2	2.3	54548	21 Z45596	DNA sequence of th
C 4	209	2.3	2116	20 X87623	Novel CD24* ORF en
C 5	209	2.3	2194	20 Z77555	Human ovarian tumo
C 6	209	2.3	2233	20 Z24431	Human bladder tumo
C 7	206.8	2.3	45546	20 X23520	Human kidney amino
C 8	205.8	2.3	9365	21 Z50359	Human CD39-L4 geno
C 9	205.4	2.3	138169	21 A34791	Human adenosine re
C 10	205.4	2.3	141589	21 A35005	Human adenosine re
C 11	205.4	2.3	141589	21 A35030	Human adenosine re
C 12	204.2	2.3	1601	21 A35191	Human adenosine re

C 13	204.2	2.3	1601	21	Z46814	Interleukin-10 (IL
C 14	204.2	2.3	1618	12	Q10207	ph15C insert conta
C 15	204.2	2.3	1618	14	Q46958	Human cytokine syn
C 16	204.2	2.3	1618	21	Z86905	Human CSIF coding
C 17	204.2	2.3	1645	21	Z58659	Human interleukin-
C 18	204.2	2.3	15630	21	A35195	Human adenosine re
C 19	203.8	2.3	7042	20	V84785	Apoptosis inducer
C 20	203.8	2.3	7042	21	Z46656	Human full-length
C 21	203.8	2.3	7075	20	V84798	Apoptosis inducer
C 22	203.6	2.3	16891	20	X37084	MEFV gene sequence
C 23	203.4	2.2	6235	21	Z29169	Human G-CSF genomi
C 24	203.4	2.2	6679	21	Z29170	Partial sequence o
C 25	203	2.2	65921	21	Z89046	Human nibrin DNA.
C 26	201.6	2.2	53526	19	T94101	Human PKD1 gene.
C 27	201.6	2.2	53577	17	T18551	Human polycystic k
C 28	201.6	2.2	53577	19	T94108	Human PKD1 locus b
C 29	200.8	2.2	1726	20	Z06248	Human secreted pro
C 30	200.6	2.2	781	20	Z15116	Human gene express
C 31	200.6	2.2	1296	19	V29031	Human protein comp
C 32	200.4	2.2	1189	20	Z00813	Human secreted pro
C 33	200	2.2	7849	16	Q94109	hML genomic DNA.
C 34	199	2.2	283	15	Q63862	AP2 sequence obtd.
C 35	199	2.2	22481	17	T11658	PEDF full length s
C 36	199	2.2	162450	21	Z86967	Retinoblastoma bin
C 37	197.2	2.2	84607	20	X90847	Human PACAP genomi
C 38	196.2	2.2	730	21	Z97408	Human prostate can
C 39	196.2	2.2	745	21	Z97407	Human prostate can
C 40	196.2	2.2	4285	20	Z25338	Human nephrin nucl
C 41	196	2.2	2867	19	V68599	Lecithin-cholester
C 42	195.8	2.2	1600	20	V69205	Sequence ID No: 11
C 43	195.8	2.2	1600	21	Z24361	Human ICAM-R leade
C 44	195.8	2.2	22481	17	T11658	PEDF full length s
C 45	195.8	2.2	106746	21	A10225	Human PCTA-1 genom

#### ALIGNMENTS

RESULT 1  
T23583  
ID T23583 standard; cDNA to mRNA; 393 BP.  
XX  
AC T23583;  
XX  
DT 02-SEP-1996 (first entry)  
XX  
DE Human gene signature HUMGS05436.

XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX WO9514772-A1.

XX PD 01-JUN-1995.

XX PF 11-NOV-1994; 94WO-JP01916.

XX PR 12-NOV-1993; 93JP-0355504.

XX PA (MATS/) MATSUBARA K.

XX PA (OKUB/) OKUBO K.

XX PI Matsubara K, Okubo K;

XX DR WPI; 1995-206931/27.

XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

XX Claim 1; Page 1404; 2245pp; Japanese.  
 PS  
 XX  
 CC A single-stranded DNA (or its complementary strand or the corres-  
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
 CC given in T19001-T26837 and which is able to hybridise to part of  
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence is unique to a particular mRNA species, almost  
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequences) as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types.  
 XX  
 XX Sequence 393 BP; 73 A; 117 C; 121 G; 78 T; 4 other;  
 SQ  
 Query Match 4.0%; Score 365.8; DB 16; Length 393;  
 Best Local Similarity 97.2%; Pred. No. 2.1e-56;  
 Matches 381; Conservative 0; Mismatches 10; Indels 1; Gaps 1;  
 1;  
 Qy 8454 gatcgagactgccccctggacctgggaacgactgagctgtcacggggttcctcctagct 8513  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 1 gatcgagactgccccctggacctgggaacgactgagctgtcacggggttcctcctagct 60  
 Qy 8514 ctccagtgaaactcctgtccagcgacacacagacgcccctatagcactgagctcacatgggact 8573  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 61 ctccagtgaaactcctgtccagcgacacacagacgcccctatagcactgagctcacatgggact 120  
 Qy 8574 gggatatgggggacatctcttccccagagagcactcagtgagctctctgtgcctggccccc 8633  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 121 gggatatgggggacatctcttccccagagagcactcagtgagctctctgtgcctggnccc 180  
 Qy 8634 agctgggccatctcttagtgagacagcttgcgcgaaactaagccagcctggctggagg 8693  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 181 agctgggccatctcttagtgagacagcttgcgcgaaactaagccagcctggctggagg 240  
 Qy 8694 agcagcagcttggggagaggatttccctgcagacctcaagccatcgcggtgggtgct 8753  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 241 agcagcagcttggggagaggatttccctgcagacctcaagccatcgcggtgggtgct 300  
 Qy 8754 gccatgacagaggctgacacctgggcagcggggctgtcacccacctctgtgcaagg 8813  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 301 gccatgacagaggctggaccnctggcagcggggtgtgtcaaccacctttgtgcaagg 360  
 Qy 8814 tggccttggac-tgcgctcaggcagagct 8844  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 361 tggccttggcttgcgctcaggcagagangt 392  
 RESULT 2  
 X30339/C  
 ID X30339 standard; DNA; 700 BP.  
 XX  
 AC X30339;  
 XX  
 DT  
 XX  
 XX 14-MAY-1999 (first entry)  
 XX  
 DE DNA encoding a human secreted protein.  
 XX  
 KW Secreted protein; cancer; tumour; neurodegenerative disorder;  
 KW developmental abnormality; foetal deficiency; blood disorder;  
 KW CNS disorder; immune system disease; autoimmune disease; hepatic disease;  
 KW renal disease; diabetes; inflammation; allergy; ischemic shock;  
 KW Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;  
 KW prostate disease; asthma; osteoporosis; arthritis; ss.  
 XX  
 OS Homo sapiens.

[illegible]





XX WPI; 1999-494537/41.  
DR P-PSDB; Y06601.  
XX  
XX New complexes of retinoblastoma protein with its binding proteins  
PT and related antibodies, nucleic acids and modulators, used for  
PT diagnosis and treatment of cancer, neurodegeneration and viral  
PT infection  
XX  
XX  
PS Claim 30; Fig 1; 103pp; English.  
XX  
CC This is the nucleotide sequence of the human CD24 gene, in which a  
CC novel open reading frame has been identified that encodes a protein  
CC (see Y06601) designated CD24\*. CD24\* was initially identified as a  
CC retinoblastoma interacting protein (RB-IP) using an improved,  
CC modified yeast two-hybrid assay system. The invention relates to  
CC complexes of RB and RB-IPs including CD24\*. The complexes modulate  
CC function of RB, i.e. cell-cycle control, cell differentiation and  
CC apoptosis, resistance to viral infection, intracellular signal  
CC transduction, transcriptional control, tumorigenesis and  
CC degeneration. Detecting abnormal levels, or function, of the  
CC RB-RB-IP complex, or levels of related RNA, is used to diagnose or  
CC monitor treatment of (pre) cancer or other hyperproliferative states  
CC such as psoriasis, neurodegeneration and viral infection. These  
CC conditions are also diagnosed by detecting levels of CD24\*, Set\*,  
CC Glu1\* or 115392\* (see Y06601-04). These disorders may be treated  
CC or prevented by administering an agent that modulates function of  
CC the RB-RB-IP complex. Animals which contain inactivated RB or  
CC RB-IP genes, or have these genes under control of non-native  
CC promoters, are used to screen for modulators of the complex.  
XX  
XX Sequence 2116 BP; 616 A; 448 C; 414 G; 638 T; 0 other;

Query Match 2.3%; Score 209; DB 20; Length 2116;  
Best Local Similarity 73.6%; Pred. No. 1.7e-28;  
Matches 282; Conservative 0; Mismatches 95; Indels 6; Gaps 1;  
QY 4275 aattattagaaagtattttataataataataacttctgtggttagcggtggttc 4334  
DB 1470 AAAAGTATTCAAACATATGCGAGAAATAAAAGCATTTTGATGGCTGGCGGGGGCTC 1411  
QY 4335 acqcttgcacccagcactttggaggcagagcgagcgagcacttgaggtcgaggaat 4394  
DB 1410 AAGCCTGTATCCACACTTGGGAGCGCGAGCGGTGGATCCTCAGGTCAGGAGT 1351  
QY 4395 tcgagatcagctggcccaatagaggaaaccccgctctctactaaaaatacaaaatcagc 4454  
DB 1350 TCTAAACAGCGCTGGGCAACATGGTGAACCCCTGTCTCTACTAAAAATTACAAAAATTAGC 1291  
QY 4455 caggcatgtgagcaggtgtctgttaataccagctacttggagggtgagcaggagaatc- 4513  
DB 1290 CGGGCATGTGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGAGGAGTATCG 1231  
QY 4514 -----agaggagagcgaggttgagtgagcagcaatcacgcactacacccagccta 4568  
DB 1230 CTTGACCCGGGAGGCAATGGTACACTGATCTGAGATCGCACCACCTGCACTCGGCCTG 1171  
QY 4569 ggtgacaaagcgagactctcaataataataataataataataataataataataata 4628  
DB 1170 GGGGCAAAAGTGAGACTGTCTAAAAATAATAATAATAATAATAATAATAATAATA 1111  
QY 4629 tgatgattattattttatcattt 4651  
DB 1110 TGGTCAATGCAATTCCTACTCTTT 1088

RESULT 5  
277555/c  
ID 277555 standard; cDNA; 2194 BP.  
XX  
AC 277555;  
XX

DT 10-APR-2000 (first entry)  
XX Human ovarian tumor cDNA library derived EST fragment 106.  
DE  
XX  
XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;  
KW gene therapy; treatment; ss.  
XX  
XX Homo sapiens.  
XX  
PN DE19817557-A1.  
XX  
PD 21-OCT-1999.  
XX  
PF 09-APR-1998; 98DE-1017557.  
XX  
PR 09-APR-1998; 98DE-1017557.  
XX  
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX  
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
XX WPI; 1999-591920/51.  
DR  
XX  
XX New nucleic acid sequences expressed in ovarian, and some other, cancer  
PT tissues, and derived polypeptides, for treatment of ovarian cancer and  
PT identification of therapeutic agents -  
PT  
XX Claim 3; Page 228-229; 310pp; German.  
PS  
XX This invention describes novel nucleic acid (cDNA) sequences (A) which  
CC have anticancer activity and are highly expressed in ovarian tumor  
CC tissue (and some also in testis and breast cancer tissue). The products  
CC of the invention can be used for gene therapy. (A) are used (i) for  
CC recombinant expression of polypeptides (B) and (ii) to isolate complete  
CC genes. (B) are used (i) to identify agents suitable for treatment of  
CC ovarian cancer; (ii) directly for treating this form of cancer (including  
CC expression from gene therapy vectors) and (iii) for generation of  
CC specific antibodies. (A) are identified by assembling ESTs (expressed  
CC sequence tags) from a particular tissue type before comparison of  
CC expression patterns. This allows a significantly longer fragment of the  
CC gene to be revealed, so should reduce the number of failures associated  
CC with the fact that ESTs from different libraries may represent different  
CC parts of the same unknown gene, distorting the estimated frequency of  
CC occurrence in a particular tissue. 277450-277572 represent the human  
CC ovarian tumor cDNA library derived EST fragments described in the method  
CC of the invention and encode the protein fragments represented in  
CC Y76505-Y76638.  
XX  
XX Sequence 2194 BP; 686 A; 431 C; 416 G; 661 T; 0 other;

Query Match 2.3%; Score 209; DB 20; Length 2194;  
Best Local Similarity 76.0%; Pred. No. 1.7e-28;  
Matches 273; Conservative 0; Mismatches 80; Indels 6; Gaps 1;  
QY 4299 aataataacttctgtggttagcggtggtgctcacgcctcaatcccgacacttgg 4358  
DB 1463 AATAAAAGCATTTTGTATGGCTGGCGCGGGCTCAAGCCTGTATCCAGCACCTTGG 1404  
QY 4359 gagcgagagcgagcagcacttgaggtcaggaatcagatcagcctggcacaatga 4418  
DB 1403 GAGGCCGAGCGGGTGTGATCACCTGAGGTGAGGTTGGAACACAGCTGGGCAACATGG 1344  
QY 4419 ggaaccccgctctactaaaaatacaaaatacagcaggtcgtggcaggtgctctga 4478  
DB 1343 TGAACCCCTGTCTCTACTAAAAATTACAAAAATTAGCCGGCATGGTGGCAGGTGCTGTA 1284  
QY 4479 atccagctacttggaggctgagcaggaatc-----agaggagcgaggttg 4532  
DB 1283 ATCCGAGCTACTTGGGAGGCTGAGGAGGAGAAATCGCTGTAGCCCGGAGGCAATGGTA 1224  
QY 4533 cagtgcagcaagatcacgccactacacccagcctcaggtgacaaagcgagactctcaaa 4592  
XX  
XX





```
XX SQ Sequence 45546 BP; 12027 A; 11359 C; 11380 G; 10780 T; 0 other;
      Query Match          2.3%; Score 206.8; DB 20; Length 45546;
      Best Local Similarity 70.3%; Pred. No. 7.5e-28;
      Matches 305; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 6855 ttctctgtttgtttggttcagggatttttttaattttatgagacaggggtctgtctgcg 6914
DB 3731 ttggctcttctttgggtttgggtttgtttgtttgttttttgagacagagctctgctctcg 3672
QY 6915 tccccaggtgagtgagtcagtgatcatatagctcactgagcctcactactcctgggc 6974
DB 3671 tcgcccagactggagtcagtcagtcacaaatcttggctcactgcaacctctgcttccctgggt 3612
QY 6975 tcaagcaatcctcctgctcctcagcctccactcaactagctgaggtacaggtgcgcgcacct 7034
DB 3611 tcaagcaattcttctgacctagacctcctcctgagtagcgggattacagggcaccgccaccat 3552
QY 7035 gcttgctaaacttttcatttttttgttagggagcgggtcctgtttttgttgcctaaagctgg 7094
DB 3551 gcttgctaa- ttttttttttttttttttttttttttttttttttttttttttttttttttt 3493
QY 7095 tctcaaaacttggtgctcgaagcaatccactgctgctgctcctccaaagtgcgtgattgc 7154
DB 3492 tctcgaaactcctgacctgacctgacctgacctgacctgacctgacctgacctgacctgacct 3433
QY 7155 agatgtgagccacgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 7213
DB 3432 agcggtcagccaccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3373
QY 7214 tgtttctgtcttttttcagaaagcagacactagcaggtgctgctgctgctgctgctgct 7273
DB 3372 tcccttctgctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 3313
QY 7274 gactgtagccacag 7287
DB 3312 ggcctcagccaaag 3299
RESULT 8
ID 250359 standard; DNA; 9365 BP.
AC 250359;
XX 18-MAY-2000 (first entry)
XX Human CD39-L4 genomic DNA.
XX CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase;
KW Arp diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;
KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW molecular weight marker; nutritional supplement; tumour; prevention;
KW drug targeting; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 72..8642
FT /*tag= a
FT /product= "Human CD39-L4 protein"
FT /note= "Coding region is interrupted with introns"
FT exon 1..288
FT /*tag= b
FT /number= 1
FT intron 289..1280
FT /*tag= c
FT /number= 1
FT exon 1281..1580
FT /*tag= d
```

```
FT intron /number= 2
FT 1581..1819 /tag= e
FT /number= 2
FT 1820..1855 /tag= f
FT /number= 3
FT 1856..2466 /tag= g
FT /number= 3
FT 2467..2555 /tag= h
FT /number= 4
FT 2556..2862 /tag= i
FT /number= 4
FT 2863..2942 /tag= j
FT /number= 5
FT 2943..3888 /tag= k
FT /number= 5
FT 3889..3950 /tag= l
FT /number= 6
FT 3951..4893 /tag= m
FT /number= 6
FT 4894..4995 /tag= n
FT /number= 7
FT 4996..5846 /tag= o
FT /number= 7
FT 5847..5987 /tag= p
FT /number= 8
FT 5988..6965 /tag= q
FT /number= 8
FT 6966..7138 /tag= r
FT /number= 9
FT 7139..8555 /tag= s
FT /number= 9
FT 8556..9365 /tag= t
FT /number= 10
XX WO200004041-A2.
XX 27-JAN-2000.
XX 16-JUL-1999; 99WO-US16180.
XX 16-JUL-1998; 98US-0118205.
XX 24-JUL-1998; 98US-0124449.
XX 04-FEB-1999; 99US-0244444.
XX 19-MAR-1999; 99US-0273447.
XX 09-JUL-1999; 99US-0350836.
XX (HYSE-) HYSEQ INC.
XX Ford J, Mulero J;
PI WPI; 2000-182397/16.
XX P-PSDB; Y44849.
XX New nucleic acid encoding human CD39-like protein, useful for treating
XX and preventing thrombotic disease -
XX Example 11; Page 112-119; 125pp; English.
PS
```







Qy	7118	atccactgccttgccctcccaagtctgagattgctgagatgtgagccacgtgcctggc	717
Db	1234	ATCACCCGCTCAGCCTCCAAAGTCTGGGATTACAGGCGTGAGCCACGCCGCCGGC	1175
Qy	7178	c 7178	
Db	1174	c 1174	
RESULT	14		
Q10207/c			
ID	Q10207	standard; cDNA; 1618 BP.	
XX	XX		
AC	Q10207;		
XX	XX		
DT	19-MAR-1991	(first entry)	
XX	XX		
DE	ph15C insert containing Human cytokine synthesis inhibitory		
DE	factor.		
XX	XX		
KW	Delayed-type hypersensitivity; DTH; leishmaniasis; parasite;		
KW	MHC-associated autoimmune disease; interferon-gamma; ds.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	33..566	
FT	FT	/*tag= a	
XX	XX		
PN	EP405980-A.		
XX	XX		
PD	02-JAN-1991.		
XX	XX		
PF	28-JUN-1990;	90EP-0307091.	
XX	XX		
PR	20-DEC-1989;	89US-0453951.	
PR	28-JUN-1989;	89US-0372667.	
XX	XX		
PA	(SCHE ) SCHERING CORP.		
XX	XX		
PI	Mosmann TR, Moore KW, Bond MW, Vieira PJM;		
XX	XX		
DR	WPI; 1991-009290/02.		
DR	R-PSDB; R10158.		
XX	XX		
PT	Mammalian cytokine synthesis inhibitory factors - capable of		
PT	inhibiting synthesis of cytokine(s) associated with delayed-type		
PT	hypersensitivity and useful in treatment of e.g. leishmaniasis		
XX	XX		
PS	Disclosure; Fig 4; 31pp; English.		
XX	XX		
CC	The gene product may be used in treatment of diseases associated		
CC	with MHC-linked immune response, suppressing a cell mediated or		
CC	humoral immune response. It may specifically be used to treat		
CC	delayed-type hyper-sensitivity, leishmaniasis, and immune disorders.		
XX	XX		
SQ	Sequence 1618 BP; 461 A; 369 C; 356 G; 432 T; 0 other;		
Query Match	2.3%;	Score 204.2; DB 12; Length 1618;	
Best Local Similarity	82.1%;	Pred. No. 1.2e-27;	
Matches	247; Conservative	0; Mismatches 53; Indels 1; Gaps	
Qy	6878	ttttattttaatttatgagacaggtctgctgtctcccccaggctggagtgcaagtgg	6937
Db	1473	TTTTATTTTATTTTGTGACAGAGTCTGCTGTCTCACCACGCTGGAGTACAGGG	1414
Qy	6938	catgatcatagctcactgagcctcatactctgggtcgaagcaatctctgctcagc	6997
Db	1413	CATGATATCAGCTCACTGCAACTTCCATCTCTGGGTTCAGCAATTCCTTGGCTCAG	1354
Qy	6998	cttcccaactagctggagctacagatgcgcgcacgcgtgcctgagctaaattttcatttt	7057

```
Db 1353 CTCCCAAGTAGCTGGATTACAGGTGCGGCCACCATGCCGCTAA-TTTTGTATTTT 1295
QY 7058 ttgtaggacggggtctctgttttgccaagctggtctcaaaactgtggcctcaagca 7117
Db 1294 TAGTAGAGACGGGGTTTACCAATTTTGACGAGCTGGTTAGGAACCTCCCTGACCTCAAGTG 1235
QY 7118 atccacctgcttggcctcccccgaagtgtgagattgcagatgtgagccaccgtgcctgac 7177
Db 1234 ATCCACCGCCTCAGCCTCCCAAGTGTCTGGATTACAGCGCTGAGCCACCGCGCCCGGC 1175
QY 7178 c 7178
Db 1174 C 1174

RESULT 15
Q46958/c
ID Q46958 standard; cDNA; 1618 BP.
XX
AC Q46958;
XX
DT 25-JAN-1994 (first entry)
XX
DE Human cytokine synthesis inhibitory factor clone ph15C.
XX
KW Mammalian cytokine synthesis inhibitory factor; CSIF; interleukin 10;
KW IL-10; immune system imbalance; human T cell; leishmaniasis;
KW rheumatoid arthritis; systemic lupus erythematosus; thyroiditis;
KW myasthenia gravis; insulin-dependent diabetes mellitus; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 31..567
FT /tag= a
FT mat_peptide 85..564
FT /tag= b
FT /product= mature_CSIF
XX
PN US5231012-A.
XX
PD 27-JUL-1993.
XX
PE 28-JUN-1989; 89US-0372667.
XX
PR 28-JUN-1989; 89US-0372667.
PR 20-DEC-1989; 89US-0453951.
PR 06-AUG-1990; 90US-0546235.
PR 20-JUL-1992; 92US-0917806.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Bond MW, Moore KW, Mosmann TR, Vieira PJM;
XX
DR WPI: 1993-249726/31.
DR P-PSDB; R39714.
XX
PT Genes and proteins encoding cytokine synthesis inhibitory factors
PT - useful in treating diseases associated with cytokine
PT imbalances, e.g. parasitic infections and auto-immune disorders
XX
PS Claim 1; Fig 4; 23pp; English.
XX
CC A human T cell cDNA library was screened with probes based on the
CC murine CSIF gene. Two clones carrying plasmids pH5C and pH15C were
CC identified. (These two expression vectors are claimed). The CSIF
CC polypeptide they encode inhibits synthesis of cytokines associated
CC with delayed type hypersensitivity responses. CSIF (also called IL-
CC 10) can be used to treat diseases associated with cytokine
CC imbalances, such as leishmaniasis and MHC-associated autoimmune
CC diseases caused by excessive production of interferon-gamma, e.g.
CC rheumatoid arthritis, SLE, IDDM, myasthenia gravis and thyroiditis.
```

```
XX
SQ Sequence 1618 BP; 463 A; 367 C; 356 G; 432 T; 0 other;

Query Match 2.3%; Score 204.2; DB 14; Length 1618;
Best Local Similarity 82.1%; Pred. No. 1.2e-27;
Matches 247; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 6878 ttttatttttaattttatgagacaggggtctgtctgtcccccaggctggagtgcaagtgg 6937
Db 1473 TTTTATTTTATTTTGTGAGACAGAGTCTTGCTGTGTCAACCAGGCTGGAGTACAGGGG 1414
QY 6938 catgatcatagctcaactgcagctcactactcctgggctcaagcaatcctcctcctcagc 6997
Db 1413 CATGATATCAGCTCACTGCAACTTCCATCTCTGGGTTTCAAGCAATTCTTTCCTCAGC 1354
QY 6998 ctccaaactagctgggactacaggtgcgcgccaccgtgctggcctaacttttcattttt 7057
Db 1353 CTCCCAAGTAGCTGGATTACAGGTGCGGCCACCATGCCGCTAA-TTTTGTATTTT 1295
QY 7058 ttgtaggacggggtctctgttttgccaagctggtctcaaaactgtggcctcaagca 7117
Db 1294 TAGTAGAGACGGGGTTTACCAATTTTGACGAGCTGGTTAGGAACCTCCCTGACCTCAAGTG 1235
QY 7118 atccacctgcttggcctcccccgaagtgtgagattgcagatgtgagccaccgtgcctgac 7177
Db 1234 ATCCACCGCCTCAGCCTCCCAAGTGTCTGGATTACAGCGCTGAGCCACCGCGCCCGGC 1175
QY 7178 c 7178
Db 1174 C 1174
```

Search completed: April 6, 2001, 00:50:49  
Job time: 17904 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:35:45 ; Search time 152.82 Seconds  
(without alignments)  
9541.804 Million cell updates/sec

Title: US-09-513-888-1

Perfect score: 9048

Sequence: 1 gctttccaagaccctgcc.....tgccatttctcaagcctct 9048

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/PCUS\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	210.6	2.3	1811	1 US-08-848-252-1	Sequence 1, Appl
C 2	209.8	2.3	35060	3 US-08-814-095-7	Sequence 7, Appl
C 3	201.6	2.2	53526	3 US-08-658-136-2	Sequence 2, Appl
C 4	201.6	2.2	53577	3 US-08-658-136-1	Sequence 1, Appl
C 5	199	2.2	1988	2 US-08-257-963B-11	Sequence 11, Appl
C 6	199	2.2	1988	4 PCT-US95-07201-11	Sequence 11, Appl
C 7	199	2.2	22481	4 PCT-US95-07201-43	Sequence 43, Appl
C 8	195.8	2.2	1600	2 US-08-487-113D-117	Sequence 117, App
C 9	195.8	2.2	1600	2 US-08-720-420A-117	Sequence 117, App
C 10	195.8	2.2	7210	2 US-08-257-963B-10	Sequence 10, Appl
C 11	195.8	2.2	7210	4 PCT-US95-07201-10	Sequence 10, Appl
C 12	195.8	2.2	22481	4 PCT-US95-07201-43	Sequence 43, Appl
C 13	195.2	2.2	7676	1 US-08-451-777A-7	Sequence 7, Appl
C 14	195.2	2.2	7676	2 US-08-451-777A-7	Sequence 7, Appl
C 15	195.2	2.2	7676	4 US-08-998-208-7	Sequence 7, Appl
C 16	195.2	2.2	7676	4 PCT-US95-06743-7	Sequence 7, Appl
C 17	194	2.1	20303	1 US-08-370-975B-6	Sequence 6, Appl
C 18	194	2.1	26764	1 US-08-370-975B-1	Sequence 1, Appl
C 19	194	2.1	87350	3 US-08-781-891-79	Sequence 79, Appl
C 20	193.8	2.1	282	1 US-08-133-629-8	Sequence 8, Appl
C 21	193.6	2.1	4823	2 US-08-457-254-5	Sequence 5, Appl
C 22	193.6	2.1	4823	2 US-08-484-257-20	Sequence 20, Appl
C 23	193.6	2.1	4823	3 US-08-999-927-5	Sequence 5, Appl
C 24	193.6	2.1	4823	4 PCT-US94-08806-28	Sequence 28, Appl
C 25	193.6	2.1	4823	4 PCT-US95-01829-5	Sequence 5, Appl
C 26	193.6	2.1	4823	4 PCT-US95-16626-5	Sequence 5, Appl
C 27	193.2	2.1	4220	1 US-08-832-883-66	Sequence 66, Appl
C 28	193.2	2.1	4220	2 US-08-832-877-66	Sequence 66, Appl

29 193.2 2.1 53526 3 US-08-658-136-2 Sequence 2, Appl  
30 193.2 2.1 53577 3 US-08-658-136-1 Sequence 1, Appl  
31 193.2 2.1 56516 2 US-08-996-306-1 Sequence 1, Appl  
C 32 193 2.1 1988 2 US-08-257-963B-11 Sequence 11, Appl  
C 33 193 2.1 1988 4 PCT-US95-07201-11 Sequence 11, Appl  
34 192 2.1 11531 1 US-08-068-945A-1 Sequence 1, Appl  
35 192 2.1 11531 1 US-07-906-806-1 Sequence 1, Appl  
C 36 191.8 2.1 17327 1 US-07-906-806-1 Sequence 15, Appl  
C 37 191.6 2.1 17327 3 US-09-128-155-17 Sequence 17, Appl  
C 38 191.4 2.1 87350 3 US-08-781-891-79 Sequence 79, Appl  
C 39 191.2 2.1 13865 3 US-09-009-217-11 Sequence 11, Appl  
40 191.2 2.1 13865 3 US-09-009-656-11 Sequence 11, Appl  
C 41 190.4 2.1 246240 2 US-08-724-394A-20 Sequence 20, Appl  
C 42 190.4 2.1 246240 2 US-08-724-394A-21 Sequence 21, Appl  
C 43 190.4 2.1 246240 2 US-08-724-394A-22 Sequence 22, Appl  
C 44 189.8 2.1 28994 3 US-08-884-324-14 Sequence 14, Appl  
45 188.8 2.1 3607 2 US-08-629-001A-8 Sequence 8, Appl

#### ALIGNMENTS

RESULT 1  
US-08-848-252-1/c  
; Sequence 1, Application US/08848252  
; Patent No. 5804177  
; GENERAL INFORMATION:  
; APPLICANT: Humphries, Keith R.  
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/848,252  
; FILING DATE: 29-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/538,052  
; FILING DATE:  
; APPLICATION NUMBER: US 08/151,672  
; FILING DATE: 15-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McDiarmid, Shona S.  
; REGISTRATION NUMBER: P-38,798  
; REFERENCE/DOCKET NUMBER: 3158-028  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; TELEX: 06-23115  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1811 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; IMMEDIATE SOURCE:  
; LIBRARY: Human  
; CLONE: Signal Transductor CD24  
; FEATURE:  
; NAME/KEY: sig\_peptide



```
FEATURE:
NAME/KEY: exon
LOCATION: complement (34528..34895)
OTHER INFORMATION: /function="arsenite resistance
OTHER INFORMATION: gene"
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: complement (34092..34358)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: complement (33779..33963)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: complement (30187..30274)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 14
FEATURE:
```

```
NAME/KEY: exon
LOCATION: complement (29945..30073)
OTHER INFORMATION: /gene="AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: complement (29664..29856)
OTHER INFORMATION: /gene="ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7

Query Match      2.3%; Score 209.8; DB 3; Length 35060;
Best Local Similarity 81.5%; Pred. No. 3.6e-35;
Matches 255; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

QY 6879 ttatttttaatttatgagacaggggtctgtctctgtccccagggctgagtgagtgagc 6938
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8387 TATTTTTTTCTTTTAAAGGACAGGGCTTGTCTCTGTACCCAGGCTGTAGTACAGTGC 8328
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6939 atgatcatgctcactgcagctcactactcctggctcagcactcctcctcctcagcc 6998
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8327 ATAATCATAGCTCAGTCAGCCCTCAAACTCCTGGGCTCAAGTCAGCTCCCACTCAGCC 8268
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6999 ttccaaactagctgggactacagdtgcgcaccgtgcctggctgaacttttctattttt 7058
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8267 TCCCCAGTAGCTGGGACTACAGGTGTGTGCCACCCCGCTGGGCTAA-TTTTTCATTTT 8209
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7059 ttagggacggggctcgtttttgtccaaaagctggtctcaaaacttggtgacctcaagcaa 7118
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8208 TGTAGAGAAAGGGGTCTTGTCTATGTGCCAGGCTGGTCTCGAACTCGTAGCTCAAGCAA 8149
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7119 tcacactgcctggcctccaaagtgcattgagattgcagatgtgagccaccgtgacctggcc 7178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8148 TCTGCCCGCTCAGCCTCCCAAAGTCTGGGATTACAGAAGTGAAGCCACTGTCTCTGGTC 8089
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7179 agattttttttt 7191
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8088 TAGCTTTTATT 8076

RESULT 3
US-08-658-136-2/c
; Sequence 2, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
```



```
/ APPLICATION NUMBER: US/08/257,963B
/ FILING DATE:
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/952,796
/ FILING DATE: 24-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DOROTHY R. AUTH
/ REGISTRATION NUMBER: 36434
/ REFERENCE/DOCKET NUMBER: 20264126051
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1988 Base Pairs
/ TYPE: Nucleic Acid
/ STRANDEDNESS: Double
/ TOPOLOGY: Unknown
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: JT108
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: 2 kb PCR product using
/ OTHER INFORMATION: primers, SEQ ID: 13 and 14.
/
/ US-08-257-963B-11

Query Match 2.2%; Score 199; DB 2; Length 1988;
Best Local Similarity 78.8%; Pred. No. 1.9e-33;
Matches 246; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 6878 tttatttttaatttatgagacagggtctgtctgtcccccaggctgagtgagtg 6937
Db 541 TTTTNNNNCTTCTGAGACGGAGTCTCGCTTGTCTNCCAAAGGCTGGAGTGCAGTGG 600

QY 6938 catgatcatagctcaactgagctgagctgagctgagctgagctgagctgagc 6997
Db 601 TGTGATCTCAGCTCACTGCAACCTCTGCCCTTGTCAAGGAATTCCTCGCTCAGC 660

QY 6998 ctccaaactagctgggactgagctgagctgagctgagctgagctgagctgagc 7117
Db 720 TAGTAGAGAGGGGGTTTACCATGTTGCGCTAGCTGGTCTCAAACTCCCGGGCTCAAGCG 779

QY 7118 atccactgcttggcctcccaagtgctgagattgagattgagccaccgtgcctggc 7177
Db 780 ATCCACCGGCTTGGGCTCCCAAAGTCTGGGATTACAGGCGTGAGCCGCGCTGGC 839

QY 7178 cagattttctt 7189
Db 840 CAGAAATAATCTT 851

RESULT 6
PCT-US95-07201-11
; Sequence 11, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIAL
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
```

```
/ STREET: 345 Park Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy Disk
/ COMPUTER: IBM PC Compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/07201
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/367,841
/ FILING DATE: 30-DEC-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/257,963
/ FILING DATE: 07-JUN-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/952,796
/ FILING DATE: 24-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DOROTHY R. AUTH
/ REGISTRATION NUMBER: 36434
/ REFERENCE/DOCKET NUMBER: 20264126PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1988 Base Pairs
/ TYPE: Nucleic Acid
/ STRANDEDNESS: Double
/ TOPOLOGY: Unknown
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: JT8A
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: 2 kb PCR product using
/ OTHER INFORMATION: primers, SEQ ID: 13 and 14; Also referred
/ OTHER INFORMATION: to as JT108
/
/ PCT-US95-07201-11

Query Match 2.2%; Score 199; DB 4; Length 1988;
Best Local Similarity 78.8%; Pred. No. 1.9e-33;
Matches 246; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 6878 tttatttttaatttatgagacagggtctgtctgtcccccaggctgagtgagtg 6937
Db 541 TTTTNNNNCTTCTGAGACGGAGTCTCGCTTGTCTNCCAAAGGCTGGAGTGCAGTGG 600

QY 6938 catgatcatagctcaactgagctgagctgagctgagctgagctgagctgagc 6997
Db 601 TGTGATCTCAGCTCACTGCAACCTCTGCCCTTGTCAAGGAATTCCTCGCTCAGC 660

QY 6998 ctccaaactagctgggactgagctgagctgagctgagctgagctgagctgagc 7057
Db 661 CTCAGAGTAGCTGGGATTACAGGCGACCTGCGCTATCATGCCACCTAA-TTTTGTATTTT 719

QY 7058 tttagggacggggtctcgtttggcgcaaaagctggtctcaaaactgtgtgctcgaagca 7117
Db 720 TAGTAGAGAGGGGGTTTACCATGTTGCGCTAGCTGGTCTCAAACTCCCGGGCTCAAGCG 779

QY 7118 atccactgcttggcctcccaagtgctgagattgagattgagccaccgtgcctggc 7177
Db 780 ATCCACCGGCTTGGGCTCCCAAAGTCTGGGATTACAGGCGTGAGCCGCGCTGGC 839
```

Qy 7178 cagatttttctt 7189  
 ||||| | |||||  
 Db 840 CAGAATAATCTT 851

## RESULT

```

PCT-US95-07201-43
: Sequence 43, Application PC/TUS9507201
: GENERAL INFORMATION:
: APPLICANT: Chader, Gerald J.; Becerra, Sofia
: APPLICANT: Patricia; Schwartz, Joan P.;
: APPLICANT: Taniwaki, Takayuki
: TITLE OF INVENTION: PIGMENT EPITHELIUM
: TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
: TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morgan & Finnegan, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy Disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/07201
: FILING DATE: 06-JUN-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/367,841
: FILING DATE: 30-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/257,963
: FILING DATE: 07-JUN-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/952,796
: FILING DATE: 24-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: DOROTHY R. AUTH
: REGISTRATION NUMBER: 36434
: REFERENCE/DOCKET NUMBER: 20264126PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 758-4800
: TELEFAX: (212) 751-6849
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 22481 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Double
: TOPOLOGY: Unknown
: MOLECULE TYPE: Genomic DNA
: FEATURE:
: NAME/KEY: pl-147
: LOCATION:
: IDENTIFICATION METHOD:
: OTHER INFORMATION: full length genomic
: OTHER INFORMATION: sequence for PEDF plus flanking sequences
: PCT-US95-07201-43

```

	Query Match	2.28;	Score 199;	DB 4;	Length 22481;
	Best Local Similarity	78.84;	Pred. No. 6e-33;		
	Matches 246;	Conservative	0;	Mismatches 65;	Indels 1; Gaps 1;
QY	6878	ttttatttttaatttatgagacagggctctgtctctgtcccccaggtcggagtcgactgg	6937		
DB	15138	TNTTTTTTTNNCTTCTGAGACGGAGTCGCTTTGTCTNCAGAGGTGGAGTGCAGTGG	15197		
QY	-6938	catgatcatagctcaactgagcctcatlactctctgggctcaagcaaatcctcctgctctcaagc	6997		

Db	15198	TG	GATCTCAGCTACATGCAACCTTGCCTCCCTGGTTCAAGGAATTC	TCTGCTGCAGC	15257	
QY	6998	cttccaactagctgggactacagtg	cgcgccacccgtgcctggctaacttttcattttt		7057	
Db	15258	CTC	CAGAGTAGCTGGGATTACAGGCAC	TGCGATCATGCCCCAGCTAA	15316	
QY	7058	tgttaggcagcgggctcgctttt	gttcaccaaacgctgtc	ctcaaaactgtgcctcaagca	7117	
Db	15317	TAGT	AGAGAGCGGGTTTACCACTGTTGCCTAGGCTGGTCTCAAACT	CCCGGGCTCAACGC	15376	
QY	7118	atccacctgccttggcctcccaaa	gctgagatgacagatg	tcagatgtagcaccgctgcctggc	7177	
Db	15377	ATCC	ACCGCGCTTGGCGCTCCCAAAGT	GCTGGGATTACAGGCGT	GAGCCACCGCGCTGGC	15436
QY	7178	cagattttttctt	7189			
Db	15437	CAGAATAATCTT	15448			

## RESULTS

US-08-487-113D-117/C  
 Sequence 117, Application US/08487113D  
 Patent No. 5837822  
 GENERAL INFORMATION:  
 APPLICANT: Gallatin, W. Michael  
 APPLICANT: Vazeux, Rosemay  
 TITLE OF INVENTION: ICAM-Related Materials and Methods  
 NUMBER OF SEQUENCES: 120  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 STREET: 6300 Sears Tower, 233 South Wacker Drive  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: United States of America  
 ZIP: 60606-6402  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/487,113D  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/286,754  
 FILING DATE: 05-AUG-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/102,852  
 FILING DATE: 05-AUG-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/009,266  
 FILING DATE: 22-JAN-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/894,061  
 FILING DATE: 05-JUN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/889,724  
 FILING DATE: 26-MAY-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/827,689  
 FILING DATE: 27-JAN-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5837822and, Greta E.  
 REGISTRATION NUMBER: 35,302  
 REFERENCE/DOCKET NUMBER: 32744  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 474-6300  
 TELEFAX: (312) 474-0448  
 TELEFAX: 25-3856  
 INFORMATION FOR SEQ ID NO: 117:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-487-113D-117

Query Match 2.2%; Score 195.8; DB 2; Length 1600;  
Best Local Similarity 77.9%; Pred. No. 8.5e-33;  
Matches 236; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 6876 gattttattttatgagacaggggtctgtctgtctgtcccccaggtgagtgagtcagt 6935  
DB 784 GGTTTTGTGTTGTTTGTGAGACAGAGTCTCACCTGTGCACCCAGGCTGGAGTGCAGT 725  
QY 6936 ggcgatgatcatgctcaactgcagcctcactctctgggtcgaagcaatcctcctgctca 6995  
DB 724 GCGGGATCTCGGCTCAGTCAGCCCTCTGCTCCCGGGTTCAACAGATTTTCTGCTCTCA 665  
QY 6996 gccctccaactagctggactcagctgcgcacccgctgctgctgaactcttcatttt 7055  
DB 664 GCTCCGAGTACTGGGATTACAGGCATGTGCCACACCGCGGCTAAATTTTGTACTT 605  
QY 7056 tttgtaggagcgggtctcgctgtttgtgccaagctggtctcaaaactgtggcctcaag 7115  
DB 604 TTTGGTAGAGACGGGGTTTCCACCATCTTGCCCAAGCTGGTCTCAAACTTGTGACCTCAAG 545  
QY 7116 caatccactgcttggcctcccaagtgatgagattgagcagattgagccacogtgcctg 7175  
DB 544 TGAGCTGCCCGCTTGGCTCCCAAGTGTGGGATTACAGGGCTGAGCCACTGCGCCCG 485  
QY 7176 gcc 7178  
DB 484 GCC 482

RESULT 9  
US-08-720-420A-117/C  
Sequence 117, Application US/08720420A  
Patent No. 5989843  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,420A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,113  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266

FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Joseph A., Jr.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 33282  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 117:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-720-420A-117

Query Match 2.2%; Score 195.8; DB 2; Length 1600;  
Best Local Similarity 77.9%; Pred. No. 8.5e-33;  
Matches 236; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
QY 6876 gattttattttatgagacaggggtctgtctgtctgtcccccaggtgagtgagtcagt 6935  
DB 784 GGTTTTGTGTTTGTGAGACAGAGTCTCACCTGTGCACCCAGGCTGGAGTGCAGT 725  
QY 6936 ggcgatgatcatgctcaactgcagcctcactctctgggtcgaagcaatcctcctgctca 6995  
DB 724 GCGGGATCTCGGCTCAGTCAGCCCTCTGCTCCCGGGTTCAACAGATTTTCTGCTCTCA 665  
QY 6996 gccctccaactagctggactcagctgcgcacccgctgctgctgaactcttcatttt 7055  
DB 664 GCTCCGAGTACTGGGATTACAGGCATGTGCCACACCGCGGCTAAATTTTGTACTT 605  
QY 7056 tttgtaggagcgggtctcgctgtttgtgccaagctggtctcaaaactgtggcctcaag 7115  
DB 604 TTTGGTAGAGACGGGGTTTCCACCATCTTGCCCAAGTGTGGGATTACAGGGCTGAGCCCTCAAG 545  
QY 7116 caatccactgcttggcctcccaagtgatgagattgagcagattgagccacogtgcctg 7175  
DB 544 TGAGCTGCCCGCTTGGCTCCCAAGTGTGGGATTACAGGGCTGAGCCACTGCGCCCG 485  
QY 7176 gcc 7178  
DB 484 GCC 482

RESULT 10  
US-08-257-963B-10/c  
Sequence 10, Application US/08257963B  
Patent No. 5840686  
GENERAL INFORMATION:  
APPLICANT: Chader, Gerald J.; Becerra, S.  
APPLICANT: Patricia; Schwartz, Joan P.;  
APPLICANT: Taniwaki, Takayuki  
TITLE OF INVENTION: PIGMENT EPITHELIUM  
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL  
TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING  
TITLE OF INVENTION: AND EXPRESSING THE PROTEIN  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan  
STREET: 345 Park Avenue

```

; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7210 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT106
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.2 kb No. 5840686 1 fragments
; OTHER INFORMATION: Derived from human placental genomic DNA
US-08-257-963B-10

```

```

Query Match      2.2%; Score 195.8; DB 2; Length 7210;
Best Local Similarity 73.7%; Pred. No. 1.7e-32;
Matches 263; Conservative 0; Mismatches 92; Indels 2; Gaps 1;

QY 6854 ttctctgtgtgtgttcagggtatttttttaatttatgagacaggtctgtctct 6913
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1893 TTTCTGTATCTGCTCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1834

QY 6914 gtcccccaggctggagtcagtgccatgatactcaactgagtgagctgagcgcacacg 6973
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1833 GTCCGCCAGGCTGGAGTGCAGTGGCTGATCTCGGTTTCTACTCAACCTCTGCTCTGGG 1774

QY 6974 ctcaagcaatctctgctcagctccaactgagtgagtgagtgagtgagtgagtgagtg 7033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1773 GTACAGCAATTAATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1714

QY 7034 tgccgtggctaaatttttttttttttttttttttttttttttttttttttttttttt 7093
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1713 CACCAGCACTAATTTTGTATTTTTCAGTAGACAGCGGGTTTTCACCATTTTGGCCAGGCTG 1654

QY 7094 gtctcaaaacttgctgcctcaagaacatcaactgctgctgctgctgctgctgctgctg 7153
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1653 GTCTTGAAGTCTGACCTCATG--ATCCACCACCTCGGCCCTCCCAAAAGTGTGGGATTA 1596

QY 7154 cagatgtgagccacagctgctgagcagatttttttttttttttttttttttttttttt 7210
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1595 CAGGCGTGAGCCACCGCGCGCGCGCGCTGGTCTTTTCTTTTCTTCTTCTTCTTCTTCT 1539

```

RESULT 11

```

PCT-US95-07201-10/c
; Sequence 10, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7210 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT6A
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.0 kb Not 1-Not
; OTHER INFORMATION: fragment; Derived from human placental
; OTHER INFORMATION: genomic DNA; also referred to as Jt106
PCT-US95-07201-10

Query Match      2.2%; Score 195.8; DB 4; Length 7210;
Best Local Similarity 73.7%; Pred. No. 1.7e-32;
Matches 263; Conservative 0; Mismatches 92; Indels 2; Gaps 1;

QY 6854 ttctctgtgtgtgttcagggtatttttttaatttatgagacaggtctgtctct 6913
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1893 TTTCTGTATCTGCTCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1834

QY 6914 gtcccccaggctggagtcagtgccatgatactcaactgagtgagctgagcgcacacg 6973
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```





```

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-451-777A-7

Query Match      2.2%; Score 195.2; DB 1; Length 7676;
Best Local Similarity 80.5%; Pred. No. 2.4e-32;
Matches 243; Conservative 0; Mismatches 53; Indels 6; Gaps 1;

QY 4312 tgtgtggttagcgtggtggtcagcctgcaatccacagcactttggaggcagagcag 4371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6063 TGAATGGGTGGCGCGGTGGCTCACACCTGTATATCCAGCACTTTGGGAAGCTGAGGCAG 6004

QY 4372 gcagagcacttgaggtcaggaattcgagatcgagcctggcccaacatgagaaacccgtct 4431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6003 GCAGATTGCTCGAGTTCAGGAGCTCGAGACCTGCGCAACATGGCGAAACCCCGTCT 5944

QY 4432 ctactaaaaatacaaaaatcagcagcaggtggtgagcaggtctgttaatccagctactt 4491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5943 CTACTAAAAATACAAAAGTTAGCCGGCATCTGCGCAGCGCTGTATATCCAGCTACTC 5884

QY 4492 ggaagcgtgagcagagagaat-----cagaggggagcggaggttcagtgagccaaga 4545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5883 AGGAGGCTGAGCGAGGAGAAATTCCTTGAACCTGGGAGCGGAGGTTGCAGTGAGTGGAGA 5824

QY 4546 tcacgccactacacccagcctagtgacaaagcagagacttctcaaatattacaataat 4605
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5823 TCATGCCACTGCACITTCAGCCTGGGTGGTGACAGAGTGAGGCTCCCTCTCAAAAAAAG 5764

QY 4606 aa 4607
   ||
Db 5763 AA 5762

RESULT 14
US-08-451-778A-7/c
; Sequence 7, Application US/08451778A
; Patent No. 5830649
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J.
; APPLICANT: Stambolian, Dwight
; TITLE OF INVENTION: Human Galactokinase Gene
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: 709 Swedeland Road/UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451.778A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10825
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Eagle, Alissa M.
; REGISTRATION NUMBER: 37,126
; REFERENCE/DOCKET NUMBER: P50268-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5364
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 7:

```



**THIS PAGE BLANK (USPTO)**

Fri Apr 6 08:40:34 2001

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database : N_Geneseq_36.*

```

RESULT	1	
T23583		
ID	T23583 standard; cDNA to mRNA; 393 BP.	
XX		
AC	T23583;	
XX		
DT	02-SEP-1996 (first entry)	
XX		
DE	Human gene signature HUMGS05436.	
XX		
KW	Gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW	human; cloning; mapping; non-biased library; diagnosis; detection;	
KW	cell typing; abnormal cell function; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9514772-A1.	
XX		
PD	01-JUN-1995.	
XX		
PF	11-NOV-1994; 94WO-JP01916.	
XX		
PR	12-NOV-1993; 93JP-0355504.	
XX		
PA	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
XX		
PI	Matsubara K, Okubo K;	
XX		
DR	WPI; 1995-206931/27.	
XX		
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	

ClaIm	1;	Page	1404;	2245pp:	Japanese.
A single-stranded DNA (or its complementary strand or the corresp. double-stranded DNA) which comprises one of the 7837 "GS" sequences given in T19001-T26837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed; the GS (gene Signature) sequences were obtained from 3'-directed CDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(rI) as the sole primer. Since the 3'- untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a CDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.					
Sequence	393 BP;	73 A;	117 C;	121 G;	78 T; 4 other;
Query Match	4.0%;	Score	365.8;	DB	16; Length 393;
Best Local Similarity	97.2%;	Pred.	No. 2.le-56;		
Matches	381;	Conservative	0;	Mismatches	10; Indels 1; Gaps
Oy	8545	gatcgactctgcccgtgcacgtggaaagcagtgcacaggccctatagcactgacctcacatggagct	8513		
Db	1	gatcgactctgcccgtgcacgtggaaagcagtgcacaggccctatagcactgacctcacatggagct	60		
Oy	8514	ctccccgaactcctgccaggcacacacagcccctatatagcactgacctcacatggagct	8573		
Db	61	ctccccgaactcctgccaggcacacacagcccctatatagcactgacctcacatggagct	120		
Oy	8574	gggatattgggggcatactcttcccagagaggcaactcagtgagctcctctgcttcggggcccc	8633		
Db	121	gggatattgggggcatactcttcccagagaggcaactcagtgagctcctctgcttcgggnccc	180		
Oy	8634	agctgggccaatctcttagttgagacagttgcccaaactaacgaccagccctgctgcagaag	8693		
Db	181	agcctggccaatctcttagttgagacagttgcccaaactaacgaccagccctgctgcagaag	240		
Oy	8694	agcagcagcttgggagaggattctccctgcagacctcaagcccatcagcggtgggtgcct	8753		
Db	241	agcagcagcttgggagaggattctccctgcagacctcaagcccatcagcggtgggtgcct	300		
Oy	8754	gccatgacagaggtgcacccctggggccagcggggtgctctcacccacctcttgtcaagg	8813		
Db	301	gccatgacagaggtgcacccctggggccagcggggtgctctcacccacctcttgtcaagg	360		
Oy	8814	tggcctttgtgc-tgcgcctgcaggcagagct	8844		
Db	361	tggcctttgtgttcgttcgcctgcaggcagangt	392		

**THIS PAGE BLANK (USPTO)**

**RESULT** 6  
**ENTRY** N21184/c  
**LOCUS** N21184 619 bp. mRNA EST 19-DEC-1995  
**DEFINITION** Y441a10.s1 Soares melanocyte 2NHMM Homo sapiens CDNA clone  
**IMAGE** IMAGE:264282 3', mRNA sequence.  
**ACCESSION** N21184  
**VERSION** N21184.1 GI:1126354  
**KEYWORDS** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**AUTHORS** 1 (bases 1 to 619)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maria, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston  
 R.K., Williamson, A., Wohlmann, P., and Wilson, R.  
**TITLE** The WashU-Merck EST Project  
**JOURNAL** Unpublished (1995)  
**COMMENT** Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 342  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 2177 Std Error: 0.00  
 Seq primer: ml3 -40 forward  
 High quality sequence stop: 342.

**THIS PAGE BLANK (USPTO)**